

# **MASTER THESIS**

**Im Studiengang  
Informationswirtschaft - Master of Science (IWM)**

## **ENTWICKLUNG EINER METHODE UND PILOTSTUDIE ZUR LANGZEITEVALUATION VON ADAPTIVEN USER INTERFACE ELEMENTEN**

**Vorgelegt von  
Kay Hauser**

**An der Fachhochschule Stuttgart - Hochschule der Medien**

**Am 05.11.2004**

**Erster Prüfer: Herr Prof. Dr. Michael Burmester**

**Zweiter Prüfer: Herr Dipl.-Inf. Klaus Wissmann  
(Siemens AG CT IC7 in München)**

Hiermit versichere ich, Kay Hauser, eidesstattlich, dass ich alle Hilfsmittel im Text und im Quellen- und Literaturverzeichnis vollständig angegeben und die Arbeit ohne fremde Hilfe selbstständig verfasst habe.

Kay Hauser

München, November 2004

# Inhaltsverzeichnis

<b>Abbildungsverzeichnis.....</b>	<b>6</b>
<b>Tabellenverzeichnis.....</b>	<b>9</b>
<b>Schlagworte.....</b>	<b>11</b>
<b>1 Einleitung.....</b>	<b>12</b>
1.1 Einführung in das Forschungsumfeld .....	12
1.2 Aufbau und Strukturierung der Arbeit .....	13
<b>2 Theorie .....</b>	<b>14</b>
2.1 Personalisierung .....	14
2.1.1 Begriffsdefinition Personalisierung .....	14
2.1.2 Formen der Personalisierung.....	14
2.2 Adaptivität.....	15
2.2.1 Begriffsdefinition Adaptivität .....	15
2.2.2 Szenarien adaptiver Systeme .....	15
2.2.3 Anforderungen an adaptive Systeme .....	17
2.2.4 Untersuchungsergebnisse zu Auswirkungen der Adaptivität auf die Usability .....	20
2.2.5 Forschungsprojekte der Hochschule der Medien Stuttgart und der Siemens AG CT IC7 zum Themenfeld Personalisierung und Adaptivität.....	22
2.3 Problemstellung .....	25
2.4 Studienfragen .....	26
2.4.1 Studienziel.....	26
2.4.2 Forschungsfragen .....	26
2.4.3 Studientyp und Gültigkeitsanspruch .....	28
2.4.4 Variablen der Studie.....	28
<b>3 Methodik .....</b>	<b>30</b>
3.1 Untersuchungsart.....	30
3.2 Erhebungsinstrumente.....	30
3.2.1 Einstellung gegenüber Formen der Personalisierung.....	30
3.2.2 Skala zur Beurteilung subjektiver mentaler Anstrengung (BSMA) zur Messung der mentalen Anstrengung einer Aufgabenbearbeitung ....	31

3.2.3	Usability-Probleme und unterstützende Effekte .....	32
3.2.4	Anforderungen an adaptive Systeme .....	34
3.2.5	Lernförderlichkeit .....	37
3.2.6	Effizienz .....	38
3.2.7	AttrakDiff zur Messung der Attraktivität und der hedonischen Qualität .....	38
3.2.8	Qualitative Daten .....	38
3.2.9	Objektive quantitative Daten .....	39
3.3	Einsatz der Erhebungsinstrumente .....	40
3.3.1	Fragebogen zum Studienbeginn .....	40
3.3.2	Feldstudie (Tagebuch) .....	42
3.3.3	Interventionen .....	44
3.3.3.1	Vor-Ort-Beobachtung .....	44
3.3.3.2	Fragebogen zur Intervention .....	45
3.3.3.3	Recall Survey .....	45
3.3.4	Dokumentation der Benutzungsoberfläche .....	46
3.3.5	Abschlussbefragung .....	46
3.3.5.1	Fragebogen zum Studienabschluss .....	46
3.3.5.2	Interview zum Studienabschluss .....	47
3.4	Durchführung der Studie .....	48
3.4.1	Auswahl des Testprodukts .....	48
3.4.2	Anwerbung und Auswahl der Studienteilnehmer .....	49
3.4.3	Einweisung der Studienteilnehmer .....	53
3.4.4	Ablauf der Studie .....	54
3.5	Quantitative Auswertung .....	55
3.6	Qualitative Auswertung .....	56
3.7	Tücken der Datenerhebung .....	57
<b>4</b>	<b>Ergebnisse .....</b>	<b>60</b>
<b>5</b>	<b>Diskussion .....</b>	<b>83</b>
<b>6</b>	<b>Ausblick .....</b>	<b>86</b>
<b>7</b>	<b>Zusammenfassung .....</b>	<b>88</b>
<b>8</b>	<b>Quellen- und Literaturverzeichnis .....</b>	<b>91</b>
<b>Anhänge .....</b>		<b>97</b>
Anhang A	Rekrutierung .....	97
Anhang B	Fragebogen zum Studienbeginn .....	98

Anhang C	Tagebuch .....	104
Anhang D	Protokoll zur Vor-Ort-Beobachtung .....	110
Anhang E	Fragebogen zur Intervention .....	113
Anhang F	Recall Survey Protokoll .....	120
Anhang G	Beispiel-Screenshots zur Dokumentation der Benutzungsoberfläche ..	121
Anhang H	Fragebogen zum Studienabschluss .....	122
Anhang I	Protokoll des Interviews zum Studienabschluss .....	125
Anhang J	Tücken der Datenerhebung: Die multitechnologische Umsetzung des BSMA im Tagebuch .....	129
Anhang K	Dritte Forschungsfrage: Ergebnisse der Erhebungsinstrumente BSMA und Anforderungen an adaptive Systeme .....	131
Anhang L	Fünfte Forschungsfrage: Auszug signifikanter Ergebnisse.....	137

## Abbildungsverzeichnis

Abbildung 1 - Szenarien adaptiver und personalisierbarer Systeme .....	16
Abbildung 2 - Erhebungsinstrument Einstellung gegenüber Formen der Personalisierung .....	30
Abbildung 3 - Erhebungsinstrument Aufgabenschwierigkeit.....	32
Abbildung 4 - Ereignisse.....	33
Abbildung 5 - Erhebungsinstrument Anforderungen an adaptive Systeme.....	35
Abbildung 6 - Erhebungsinstrument Anforderungen an adaptive Systeme Abschlussbefragung Beispiel Dimension Simplicity.....	36
Abbildung 7 - Erhebungsinstrument Lernförderlichkeit.....	37
Abbildung 8 - Erhebungsinstrument Effizienz .....	38
Abbildung 9 - Erste Forschungsfrage: Ergebnis Parameter Anzahl „Usability-Probleme“ über den Studienzeitraum .....	61
Abbildung 10 - Erste Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Usability-Probleme“ über den Studienzeitraum .....	61
Abbildung 11 - Erste Forschungsfrage: Ergebnis Parameter Intensität „Usability-Probleme“ über den Studienzeitraum .....	62
Abbildung 12 - Erste Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Unterstützende Effekte“ über den Studienzeitraum .....	62
Abbildung 13 - Erste Forschungsfrage: Ergebnis Parameter Intensität „Unterstützende Effekte“ über den Studienzeitraum .....	63
Abbildung 14 - Erste Forschungsfrage: Ergebnis Parameter Anzahl „Usability-Probleme“ nach Phasen adaptiver Anpassung, Adaptive Gruppe .....	65
Abbildung 15 - Erste Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Usability-Probleme“ nach Phasen adaptiver Anpassung, Adaptive Gruppe .....	66
Abbildung 16 - Erste Forschungsfrage: Ergebnis Parameter Intensität „Usability-Probleme“ nach Phasen adaptiver Anpassung, Adaptive Gruppe .....	66

Abbildung 17 - Zweite Forschungsfrage: Ergebnis Lernförderlichkeit - Zeit .....	69
Abbildung 18 - Zweite Forschungsfrage: Ergebnis Lernförderlichkeit - Exploration....	70
Abbildung 19 - Zweite Forschungsfrage: Ergebnis Lernförderlichkeit - Menge und Details .....	70
Abbildung 20 - Zweite Forschungsfrage: Ergebnis Lernförderlichkeit - Einprägbarkeit .....	71
Abbildung 21 - Dritte Forschungsfrage: Ergebnis Effizienz .....	73
Abbildung 22 - Dritte Forschungsfrage: Ergebnis Abschlussfragen Anforderungen an adaptive Systeme .....	75
Abbildung 23 - Rekrutierung (Dateianhang, Aushang, Flugzettel) .....	97
Abbildung 24 - Screenshot Dokumentation der Benutzeroberfläche Beispiel 1 .....	121
Abbildung 25 - Screenshot Dokumentation der Benutzeroberfläche Beispiel 2 .....	121
Abbildung 26 - Dritte Forschungsfrage: Ergebnis BSMA Tagebuch.....	131
Abbildung 27 - Dritte Forschungsfrage: Ergebnis BSMA Fragebogen.....	131
Abbildung 28 - Dritte Forschungsfrage: Ergebnis Etiquette Tagebuch.....	132
Abbildung 29 - Dritte Forschungsfrage: Ergebnis Etiquette Fragebogen.....	132
Abbildung 30 - Dritte Forschungsfrage: Ergebnis Privatsphäre Tagebuch .....	132
Abbildung 31 - Dritte Forschungsfrage: Ergebnis Privatsphäre Fragebogen .....	133
Abbildung 32 - Dritte Forschungsfrage: Ergebnis Verständnis Tagebuch .....	133
Abbildung 33 - Dritte Forschungsfrage: Ergebnis Verständnis Fragebogen .....	133
Abbildung 34 - Dritte Forschungsfrage: Ergebnis Vertrauen Tagebuch .....	134
Abbildung 35 - Dritte Forschungsfrage: Ergebnis Vertrauen Fragebogen .....	134
Abbildung 36 - Dritte Forschungsfrage: Ergebnis Breadth of Experience Tagebuch ..	134
Abbildung 37 - Dritte Forschungsfrage: Ergebnis Breadth of Experience Fragebogen .....	135

Abbildung 38 - Dritte Forschungsfrage: Ergebnis Kontrolle Tagebuch.....	135
Abbildung 39 - Dritte Forschungsfrage: Ergebnis Kontrolle Fragebogen.....	135
Abbildung 40 - Dritte Forschungsfrage: Ergebnis Nutzen Tagebuch.....	136
Abbildung 41 - Dritte Forschungsfrage: Ergebnis Nutzen Fragebogen.....	136
Abbildung 42 - Fünfte Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Unterstützende Effekte" über den Studienzeitraum .....	137
Abbildung 43 - Fünfte Forschungsfrage: Ergebnis Parameter Intensität „Unterstützende Effekte" über den Studienzeitraum.....	137
Abbildung 44 - Fünfte Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Usability-Probleme" über den Studienzeitraum .....	138
Abbildung 45 - Fünfte Forschungsfrage: Ergebnis Parameter Intensität „Usability-Probleme" über den Studienzeitraum .....	138
Abbildung 46 - Fünfte Forschungsfrage: Ergebnis BSMA Tagebuch.....	139
Abbildung 47 - Fünfte Forschungsfrage: Ergebnis BSMA Fragebogen.....	139
Abbildung 48 - Fünfte Forschungsfrage: Ergebnis Breadth of Experience Tagebuch .	140
Abbildung 49 - Fünfte Forschungsfrage: Ergebnis Nutzen Tagebuch.....	140
Abbildung 50 - Fünfte Forschungsfrage: Ergebnis Verständnis Tagebuch .....	141
Abbildung 51 - Fünfte Forschungsfrage: Ergebnis Etiquette Tagebuch.....	141
Abbildung 52 - Fünfte Forschungsfrage: Ergebnis Kontrolle Tagebuch.....	142



## **Tabellenverzeichnis**

Tabelle 1 - Matrix der Studienbewerber .....	52
Tabelle 2 - Matrix der Studienteilnehmer .....	52
Tabelle 3 - Matrix der Testgruppen .....	53
Tabelle 4 - Matrix der Studienteilnehmer aufgrund der Auswertung der qualitativen Daten .....	58
Tabelle 5 - Stichprobenumfang (Anzahl Tagebucheinträge) erste Forschungsfrage: Betrachtung des Nutzungszeitraums .....	60
Tabelle 6 - Stichprobenumfang (Anzahl Tagebucheinträge) erste Forschungsfrage: Betrachtung Phasen adaptiver Anpassungen bei der Adaptiven Gruppe .....	64
Tabelle 7 - Erste Forschungsfrage: Ergebnis Wahrscheinlichkeit einzelner „Usability-Probleme“ bei der Adaptiven Gruppen über den Studienzeitraum .....	67
Tabelle 8 - Erste Forschungsfrage: Ergebnis Wahrscheinlichkeit einzelner „Usability-Probleme“ bei der Adaptiven Gruppe nach Phasen adaptiver Anpassungen .....	67
Tabelle 9 - Erste Forschungsfrage: Ergebnis objektiver quantitativer Daten .....	68
Tabelle 10 - Stichprobenumfang (Interventionen) zweite Forschungsfrage.....	69
Tabelle 11 - Stichprobenumfang (Tagebucheinträge) dritte Forschungsfrage: Datenerhebung Tagebuch.....	72
Tabelle 12 - Stichprobenumfang (Interventionen) dritte Forschungsfrage: Datenerhebung Fragebogen.....	72
Tabelle 13 - Stichprobenumfang (Abschlussbefragungen) dritte Forschungsfrage: Datenerhebung Fragebogen zum Studienabschluss .....	72
Tabelle 14 - Stichprobenumfang (Interventionen) vierte Forschungsfrage .....	77
Tabelle 15 - Vierte Forschungsfrage: Vergleich der Dimensionen des AttrakDiff .....	77
Tabelle 16 - Stichprobenumfang (Tagebucheinträge) fünfte Forschungsfrage : Datenerhebung Tagebuch .....	78

Tabelle 17 - Stichprobenumfang (Interventionen) fünfte Forschungsfrage: Datenerhebung Fragebogen .....	79
Tabelle 18 - Ergebnisse zur Evaluation der Erhebungsinstrumente .....	82

## **Schlagworte**

User-Centered Design ; User Interface Design

Gebrauchstauglichkeit ; Usability ; Ease of Use

Personalisierung ; Adaptivität

Langzeitevaluation ; Tagebuchstudie ; Feldstudie ; Fragenbogenstudie

Microsoft Word XP

# 1 Einleitung

## 1.1 Einführung in das Forschungsumfeld

Die vorliegende Ausarbeitung wurde bei der Siemens AG Corporate Technology im Fachzentrum für User Interface Design (CT IC7) angefertigt. Das Fachzentrum CT IC7 beschäftigt sich mit der Entwicklung, Gestaltung und Implementierung von *User Interfaces* (im weiteren Verlauf als UI bezeichnet)<sup>1</sup>.

Der entsprechende methodische Ansatz findet sich im *User-Centered Design* (im weiteren Verlauf als UCD bezeichnet) wieder, welches die benutzerorientierte Gestaltung von Produkten zum Ziel hat. Beim UCD stehen die Bedürfnisse und Anforderungen der Menschen an die Technik im Vordergrund. Systeme sollen so gestaltet sein, dass der Nutzer sie einfach und problemlos bedienen kann und er bei seinen Arbeitsaufgaben effizient unterstützt wird (DIN EN ISO 13407, 1999).

Ziel jedes Konzepts zur Gestaltung einer Benutzungsoberfläche ist die Gebrauchstauglichkeit des Produkts, in der Literatur spricht man auch von der *Usability* oder dem *Ease of Use* eines Produkts. Die ISO-Norm 9241-11 (DIN EN ISO 9241-11, 1998), in der die Anforderungen an die Gebrauchstauglichkeit eines Produkts spezifiziert sind, definiert diese wie folgt:

„Gebrauchstauglichkeit ist das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und zufriedenstellend zu erreichen.“ (DIN EN ISO-Norm 9241-11, 1998, S.4)

Unter dem Nutzungskontext werden die Benutzer selbst, die Arbeitsaufgaben, die Arbeitsmittel sowie die physische und soziale Umgebung verstanden. Die Effektivität ist die Genauigkeit und Vollständigkeit, die Effizienz ist der eingesetzte Aufwand, mit dem Benutzer ein Ziel erreichen. Die Zufriedenstellung beschreibt die Freiheit von Beeinträchtigungen und positive Einstellungen gegenüber der Produktnutzung.

Um die Gebrauchstauglichkeit eines Produkts zu erreichen, spezifiziert die ISO-Norm 9241-10

---

<sup>1</sup> Ein User Interface (deutsch: Benutzungsoberfläche) bezeichnet die Aspekte eines Systems, die ein Mensch verwendet, um mit dem System zu interagieren (Preece, 1993). Beispiele sind ein Bildschirm oder eine Tastatur, aber auch ein An/Aus-Schalter oder eine Spracherkennung.

in ihren „Grundsätzen der Dialoggestaltung“ die folgenden Richtlinien (DIN EN ISO-Norm 9241-10, 1996) für die Gestaltung und Bewertung:

- Aufgabenangemessenheit,
- Selbstbeschreibungsfähigkeit,
- Steuerbarkeit,
- Erwartungskonformität,
- Fehlertoleranz,
- Lernförderlichkeit,
- Individualisierbarkeit.

Der Grundsatz zur Individualisierbarkeit stellt die Notwendigkeit heraus, dass ein System Anpassungen an Erfordernisse der Arbeitsaufgabe, Vorlieben und Fähigkeiten des Nutzers zulässt. Diese Beschreibung enthält die gleichen Ansätze wie das in der vorliegenden Arbeit untersuchte Forschungsfeld der Personalisierung, welches sich mit der Anpassung von Produkten an die Wünsche und Anforderungen eines Nutzers beschäftigt. Der Studienfokus liegt auf der, als Adaptivität bezeichneten, vom System initiierten Form der Personalisierung und deren Auswirkungen auf die Gebrauchstauglichkeit eines Produkts.

## ***1.2 Aufbau und Strukturierung der Arbeit***

In Kapitel 2, dem Theorieteil, wird das Forschungsfeld Personalisierung und Adaptivität ausführlich erläutert und der bisherige Forschungsstand dargestellt. Darauf basierend folgt die Ableitung der Forschungsfragen.

Der folgende Methodikteil in Kapitel 3 beschreibt die eingesetzten Erhebungsinstrumente, die Untersuchungsobjekte sowie die Durchführung der Studie.

Die Darstellung der Untersuchungsergebnisse, entsprechend der Forschungsfragen, erfolgt in Kapitel 4.

Ergebnisinterpretation und Diskussion bilden Kapitel 5.

Kapitel 6 beschließt die Studie mit einem Ausblick auf zukünftige Forschungsansätze und weitere Entwicklungen des Forschungsfeldes.

## **2 Theorie**

### **2.1 Personalisierung**

#### **2.1.1 Begriffsdefinition Personalisierung**

*Personalisierung* beschreibt eine Reihe von Technologien und Funktionen, die im Bereich der Endnutzung eines Produkts eingesetzt werden (Kramer, Noronha, Vergo 2000). Durch Anpassungen an individuelle Wünsche und Anforderungen soll der Nutzer seine Ziele im Umgang mit dem Produkt erreichen. Personalisierung beinhaltet die Anpassungen von Prozessen, Funktions- und Interaktionsabläufen, Informationsinhalten, Benutzungsoberflächen oder äußeren Merkmalen (Blom, 2000). Die Spanne der Einsatzmöglichkeiten reicht von der Darstellung des Nutzernamens bis zur Konfiguration von Produktfunktionen.

Vom Nutzer aus kann die Motivation zur Personalisierung durch den Wunsch zur Arbeitserleichterung entstehen, wie z.B. dem einfachen Zugang zu Information, der Anpassung an Arbeitsziele, -gewohnheiten und -gegebenheiten, oder durch soziale Bedürfnisse, wie beispielsweise Emotionen und dem Wunsch der Persönlichkeitsdarstellung (Blom, 2000). Als wichtigste Gründe für die Personalisierung ermittelte Blom eine funktionsorientierte Gestaltung des personalisierbaren Systems, einer Entscheidungsfindung dienliche Ergebnispräsentation, die Notwendigkeit zusätzliche Informationen zu erhalten, die Systemkontrolle und das Wissen über die Personalisierungsmethodik (Blom, 2002).

Aus Sicht eines Unternehmens oder Produkthanbieters hat Personalisierung das Ziel eine stärkere Bindung des Nutzers zu dem Produkt aufzubauen und somit die Loyalität des Kunden zu erhöhen (Riecken, 2000).

#### **2.1.2 Formen der Personalisierung**

Personalisierung lässt sich anhand der Steuerung der Anpassungen in zwei Arten gliedern.

Die *Adaptierbarkeit* beschreibt eine manuelle, durch den Nutzer initiierte Personalisierung. Der Nutzer kann dabei zwischen verschiedenen Gestaltungs- und Funktionsvarianten wählen, die ihm das System zur Verfügung stellt (Nitschke, 2001). Synonym verwendete Begriff für Adaptierbarkeit sind Adaption, Customization, Individualisierbarkeit und manuelle Personalisierung.

Die *Adaptivität* beschreibt hingegen eine vom System gesteuerte, automatische Personalisierung.

Beide Formen der Personalisierung können innerhalb eines Systems unterstützend eingesetzt werden (Blom, 2000), diese nennt man *Mixed-Initiative UIs*.

## **2.2    *Adaptivität***

### **2.2.1    Begriffsdefinition Adaptivität**

Langley (Langley, 1997) beschreibt Adaptivität als ein interaktives System, welches mit dem Nutzer auf der Basis von gesammelten Erfahrungen interagiert und sich dessen Gewohnheiten anpasst:

“An adaptive user interface is an interactive software system that improves its ability to interact with a user based on partial experience with the user.“ (Langley, 1997, zit. nach Preim, 1999, S.465)

Um sich den Gewohnheiten des Nutzers anpassen zu können, muss ein adaptives System, vor allem in der englischsprachigen Literatur auch oft als Agent UI bezeichnet, aufgrund von Interaktionsdaten Präferenzen und Ziele des Benutzers erschließen, sammeln und ein Modell zu Annahmen über das Nutzerverhalten aufbauen. Dieses Modell muss vom System laufend aktualisiert und auf die Gültigkeit der Annahmen hin überprüft und angepasst werden (Kobsa, 1993). Das Nutzermodell dient dem System dann als Datenbasis für die Initiierung von Anpassungen und Interaktionen.

### **2.2.2    Szenarien adaptiver Systeme**

Szenarien für adaptive Systeme sind sehr weitreichend und können unterschiedlichste Formen annehmen wie

- die Anpassung der Menüs bei Office Anwendungen,
- die Anpassung der Symbolleisten bei Office Anwendungen,
- eine Liste zuletzt geöffneter Programme oder Dateien bei Windows Betriebssystemen,
- eine Liste getätigter Anrufe bei Mobiltelefonen,
- adaptive Hypermediasysteme, z.B. Anpassung an die Nutzerterminologie bei Webseiten,

- Location Based Services, z.B. Informationen zu Reiseorten bei Navigationssystemen,
- Recommender Systeme, z.B. Buchvorschläge bei der Webseite Amazon.de.

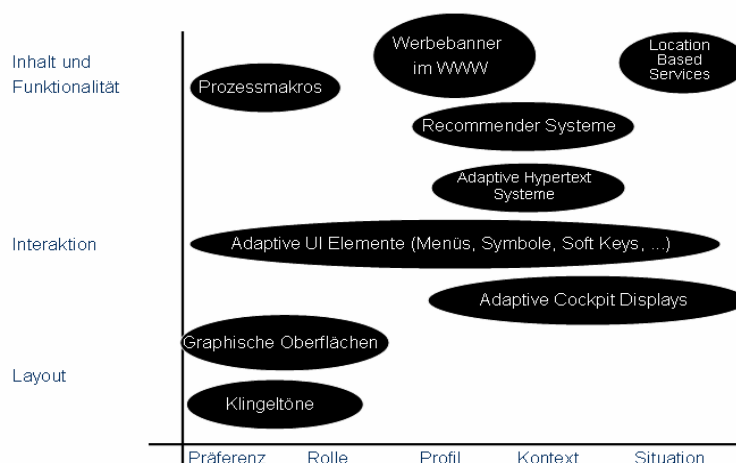
Um adaptive und personalisierbare Systeme in den Gestaltungsprozess eines UI einbeziehen zu können, ist eine Strukturierung der verschiedenen Szenarien sinnvoll. Eine Unterscheidung erfolgt bzgl. des Anpassungsobjekts und der Motivation zur Anpassung (Blom, 2000). Anpassungsobjekte können in drei Ausprägungen beschrieben werden:

- Inhalt und Funktionalität als das Ergebnis, das ein Anwender mit der Nutzung eines Systems verbindet,
- Interaktion zwischen Nutzer und System,
- Layout des UI.

Die Motivation lässt sich in fünf Kategorien unterteilen:

- Präferenzen des Nutzers,
- Rolle des Nutzers gegenüber dem System wie etwa Anwender oder Administrator,
- Profil des Nutzers,
- Kontext der Nutzung,
- Situation, in der das System genutzt wird.

Abbildung 1 stellt beide Merkmale mit ihren Ausprägungen in einer Matrix dar und illustriert diese anhand einiger Beispielszenarien:



**Abbildung 1 - Szenarien adaptiver und personalisierbarer Systeme (Wissmann, 2003)**



### 2.2.3 Anforderungen an adaptive Systeme

Aufgrund der für adaptive Systeme typischen Charakteristik eines interagierenden Systems müssen an die Gestaltung des UI spezielle Anforderungen gestellt werden, damit die Usability des Produkts gewährleistet wird. Einige Anforderungen korrespondieren mit allgemeinen Usability Richtlinien, wie etwa den Dialogprinzipien der ISO-Norm 9241-10 (DIN EN ISO 9241-10, 1996), manche Anforderungen gehen aber über die Prinzipien „normaler“ Systeme hinaus und sind speziell für die Gestaltung adaptiver Systeme relevant.

#### Nutzen

Horvitz stellt generell die Forderung auf, dass für den Nutzer durch vom System automatisierte Aufgaben und adaptive Anpassungen ein zusätzlicher Nutzen oder Mehrwert entstehen muss (Horvitz, 1999). Der zusätzliche Nutzen eines adaptiven Systems kann nach Jameson erfolgen durch (Jameson, 2003)

- die Übernahme von Routineaufgaben durch das System,
- die Anpassung der Benutzungsoberfläche,
- Hilfestellungen bei der Nutzung des Systems,
- eine Dialogführung für den Nutzer,
- Hilfen bei der Informationssuche,
- die Aufbereitung von Information und die Informationspräsentation,
- Recommender Systeme, die beispielsweise Kaufvorschläge entsprechend des Nutzerprofils abgeben,
- Supporting Collaboration zur Empfehlung möglicher Partnernutzer entsprechend vorhandener Nutzerprofile,
- die Verbesserung der Lernförderlichkeit.

#### Breadth of Experience

Wichtige Faktoren zum Erreichen eines Nutzens, durch eine vom System initiierte Interaktion, sind stichhaltige Annahmen zum Nutzerverhalten und die Fähigkeit diese auch in den richtigen Momenten einzusetzen. Die Anforderungen an adaptive Systeme bestehen im Optimieren der Bestimmung von Nutzerintentionen, im Timing für Systeminitiativen und in den Inhalten einer Interaktion sowie in der Minimierung von Folgen falscher Systeminitiativen (Horvitz, 1999). Jameson nennt diese Optimierungen „Breadth of Experience“ und meint eine

„Anpassungsgeduld“ des adaptiven Systems, Interaktionen nicht auf der Basis eines unvollständigen Nutzermodells zu initiieren (Jameson, 2003).

### Vertrauen

Aufgrund der automatisierten Anpassungen und der Übernahme von Aufgaben besteht die Notwendigkeit adaptive Systeme so zu gestalten, dass der Nutzer dem System zutraut diese auch entsprechend seiner Erwartungen ausführen zu können (Wexelblatt, Maes, 2004). Das Aufbauen einer Vertrauensbasis entsteht vor allem durch Erfahrungen mit der Funktionalität und Zuverlässigkeit des Systems (Parasuraman, Miller, 2004).

### Verständnis

Shneiderman (Shneiderman, 1995) stellt heraus, dass die Akzeptanz eines adaptiven System auch entscheidend davon abhängt, ob ein Nutzer weiß, was das System macht und ob er Instruktionen und Systemmitteilungen versteht.

„If users are unsure about what the agent will do, if they cannot repeat previously successful actions, or if they cannot understand instructions, dialog boxes, and error messages, they will not tolerate the interface.“ (Shneiderman, 1995, S.13)

Jameson beschreibt diese Anforderungen detaillierter als die Vorhersagbarkeit von Effekten einer vom Nutzer initiierten Interaktion und die Transparenz des Systemverhaltens. Diese wirken auf den Ebenen Layout und Antwortverhalten eines Systems, Erfolg einer Interaktion und auf der übergeordneten Nutzungsumgebung (Jameson, 2003). Layout und Antwortverhalten beziehen sich dabei auf die gestalterische Umsetzung von UI-Elementen. Ein vorhersagbares Aussehen von Symbolen oder Menüs ermöglicht den Nutzern einen schnellen Zugriff auf Funktionen und eine einfache Handhabung der UI-Elemente. Der Erfolg einer Interaktion beschreibt z.B. das Resultat einer Datenbanksuche. Dabei möchte der Nutzer nicht jeden Arbeitsschritt des Systems verstehen, sondern nachvollziehen, wie er die Qualität des Endergebnisses beeinflussen kann. Die Vorhersagbarkeit und Transparenz auf der übergeordneten Nutzungsumgebung umfassen global alle Verhaltensweisen eines Systems.

### Kontrollierbarkeit

In einem engen Zusammenhang mit dem Systemverständnis sieht Jameson die Kontrollierbarkeit des Systems und meint die Möglichkeit für den Nutzer, eine Interaktion oder einen Systemzustand herbeizuführen oder verhindern zu können (Jameson, 2003). Wexelblatt

und Maes sehen die Kontrollierbarkeit des Systems im Wechselspiel mit dem Vertrauen in das System. Vertraut ein Nutzer dem adaptiven System und den Anpassungen, die es vornimmt, ist er auch bereit seine Kontrolle abzugeben und somit den Nutzen des Systems auszuschöpfen. Um beides zu gewährleisten, schlagen sie die Möglichkeit vor, die Autonomie des Systems und somit auch die Kontrollierbarkeit in variablen Levels einstellen zu können (Wexelblatt, Maes, 2004).

Lanier verdeutlicht, durch einen düsteren Ausblick auf zukünftige Entwicklungen von Agent UIs, die Wichtigkeit der Kontrollierbarkeit von adaptiven Systemen. Er sieht einen völligen Kontrollverlust des Nutzers aufgrund einer Verselbstständigung der Systeme voraus, die sich nicht dem Nutzer anpassen, sondern denen sich der Nutzer anpassen muss (Lanier, 1995).

### Privatsphäre

Eine weitere soziale Komponente der Kontrollierbarkeit besteht in der Wahrung der Privatsphäre des Nutzers (Norman, 1994). Ausgelöst durch die Speicherung von individuellen Nutzerdaten zu Interaktionen bis hin zu finanziellen Transaktionen durch das adaptive System können beim Nutzer Bedenken entstehen, ob das System diese Daten nicht für ungeeignete Zwecke verwendet (Jameson, 2003).

### Etiquette

Parasuraman und Miller beschreiben die Anforderung einer Etiquette als das Benehmen des adaptiven Systems gegenüber dem Nutzer bei der Darstellung vom System initiierten Interaktionen.

„We define Etiquette... as the set of prescribed and proscribed behaviours that permits meaning and intent to be ascribed to actions.” (Parasuraman, Miller, 2004, S. 52)

Ein angemessener Kommunikationsstil des Systems kann dabei, ähnlich wie bei der menschlichen Kommunikation zu einer sozialen Anziehung, zu dem Entstehen von Vertrauen in das System führen (Parasuraman, Miller, 2004). Somit besteht die Notwendigkeit vom System ausgehende Hinweise, Nachfragen oder Aufforderungen so zu gestalten, dass der Nutzer möglichst wenig bei der Ausführung seiner eigentlichen Aufgabe gestört wird (Jameson, 2003). Empfehlenswert ist es dabei, die Kommunikation in der Sprache des Nutzers und nicht in der Sprache des Systems zu verfassen (Wexelblatt, Maes, 2004).

### Integration und Zugriff

Der Nutzer sollte die Möglichkeit haben, auf das Modell des adaptiven Systems zugreifen zu können. Dafür sollte das Agent UI so weit wie möglich in das Gesamtsystem integriert sein und darüber hinaus einfache Konfigurationsmöglichkeiten bieten (Wexelblatt, Maes, 2004). Der Ansatz einer Konfigurationsmöglichkeit von adaptiven Systemen entspricht der Horvitzschen Idee von Mixed-Initiative UIs und birgt neben Synergiepotentialen aber auch das Risiko, dass Systeme zu komplex werden (Horvitz, 1999). Die Konfigurationsmöglichkeiten sollten so gestaltet sein, dass der Nutzer keine neuen Funktionen für die Personalisierung lernen muss (Wexelblatt, Maes, 2004).

#### **2.2.4 Untersuchungsergebnisse zu Auswirkungen der Adaptivität auf die Usability**

In der Fachliteratur findet man wenige Studien, die Auswirkungen adaptiver oder allgemein personalisierbarer Systeme auf die Usability eines Produkts hin untersucht haben. Ergebnisse vorhandener Untersuchungen sollen im folgenden dargestellt werden.

Findlater und McGrenere führten einen Usability-Test zum Vergleich der Effizienz statischer, adaptierbarer und adaptiver Menüs mit Prototypen<sup>2</sup> von Microsoft Word 2000 durch (Findlater, McGrenere, 2004). Gemessen wurden sowohl objektive quantitative Daten, wie die Geschwindigkeit der Aufgabenbearbeitung und benötigte Aufgabenschritte als auch subjektive Einschätzungen der Probanden zur Präferenz eines Systems, Effizienz, Fehlerrate, Frustration und Ease of Use mittels eines Fragebogens. Insgesamt sprach sich die Mehrheit der Studienteilnehmer für personalisierbare UIs aus, den stärksten Zuspruch fanden die adaptierbaren Menüs. Die objektiven Messwerte ergaben jedoch, dass die Studienteilnehmer ihre Aufgaben am effizientesten mit den statischen Menüs ausführten.

Höök untersuchte in einem Usability-Test zwei Varianten eines Hypermediasystems, einmal mit adaptiver und einmal mit nicht-adaptiver Funktionalität (Höök, 1997). Während des Tests wurden die Probanden gefilmt und ihre Interaktionen mittels eines Softwaretools dokumentiert, welches Daten zur Geschwindigkeit der Aufgabenbearbeitung, Bearbeitungsschritten, etc. lieferte. Zum Abschluss des Tests wurden die Teilnehmer zu ihren Eindrücken befragt. Die

---

<sup>2</sup> Ein Prototyp ist ein Arbeitsmodell eines endgültigen Produkts (Weiss, 2002). Er simuliert das UI und wichtige Funktionen und dient zum Sammeln von ersten Eindrücken und Erfahrungen.

Studie ergab, dass die Probanden das adaptive Hypermediasystem gegenüber dem nicht-adaptiven bevorzugten, sie mit den Anpassungen des adaptiven Systems zufrieden waren und die an sie gestellten Aufgaben auch in weniger Bearbeitungsschritten lösen konnten.

In einer Feldstudie<sup>3</sup> untersuchten McGrenere, Baecker und Booth (McGrenere, Baecker, Booth, 2002) die Unterschiede zwischen einem adaptierbaren UI-Prototypen von Microsoft Word und Microsoft Word 2000. Die Probanden konnten während der Untersuchung wahlweise zwischen den beiden Systemen hin und her schalten. Für die Studie wurden die Teilnehmer anhand der von McGrenere und Moore entworfenen Variable Feature-keen/Feature-shy (McGrenere, Moore, 2000) nach ihren individuellen Vorlieben bezüglich des Funktionsumfangs, der Vollständigkeit und der Aktualität eines Systems in zwei Testgruppen unterteilt. Über einen Studienzeitraum von sechs Wochen mussten die Probanden mehrere Aufgaben während insgesamt drei Untersuchungsterminen lösen. Für bestimmte Aufgaben wurde bei Microsoft Word 2000 die Adaptivität der Menüs aktiviert bzw. deaktiviert, was den Probanden aber nicht mitgeteilt wurde. Anhand von Log-Dateien wurden die Nutzungszeiten jeder Systemvariante für die Aufgabenbearbeitung festgehalten, durch Fragebogen und Interviews die Einschätzungen bezüglich der Zufriedenheit, Navigierbarkeit, Kontrollierbarkeit und Erlernbarkeit der Systeme ermittelt. Ergebnisse der Studie zeigten, dass die Nutzer die meiste Zeit mit dem adaptierbaren Prototypen arbeiteten, diesen auch als navigierbarer, kontrollierbarer und erlernbarer bewerteten. Allerdings bemerkte ein Drittel der Probanden nicht die adaptive Funktionalität von Microsoft Word 2000 und bewertete dieses System dann auch schlechter.

In ihrer Diplom-Arbeit untersuchte Dubiella (Dubiella, 2003) die Anforderungen an personalisierbare Handys und deren Akzeptanz bei den Nutzern anhand von Focus Groups<sup>4</sup> und einer Fragebogenstudie.

Aus den drei Nutzergruppen, professionelle Nutzer, Jugendliche und Senioren, wurden Probanden für die Focus Groups rekrutiert. Das erste Ziel bestand im Sammeln von Ideen und Wünschen zur Personalisierung von Handys. In einem weiteren Schritt bekamen die Teilnehmer

---

<sup>3</sup> Feldstudien untersuchen Probanden in ihrer natürlichen Arbeitsumgebung wie ihrem Arbeitsplatz, zu Hause oder an ihrer Schule. Ziel ist es so mehr über die normalen Arbeitsgewohnheiten, das Verhalten und Ziele der Nutzer zu erfahren (Kantner, Sova, Rosenbaum, 2003).

<sup>4</sup> Eine Focus Group ist eine organisierte, durch einen Moderator gesteuerte Diskussion mit einem ausgewählten Personenkreis, entsprechend einer spezifischen Nutzergruppe. Das Ziel besteht im Ermitteln von Ideen, Wünschen und Erfahrungen zu einem bestimmten Thema oder Produkt (Gibbs, 1997).

zur Bewertung einen Software-Prototypen vorgestellt, mit dem sie ihr Handy am PC personalisieren konnten. Außerdem wurde anhand von Beispielen eine adaptive Form der Personalisierung illustriert.

Die Resultate der Focus Groups flossen in einen online verfügbaren Fragebogen ein. Die Fragen galten ebenfalls den Wünschen und der Akzeptanz der Nutzer an personalisierbare Handys sowie deren Präferenzen bezüglich einer Form der Personalisierung und Anforderungen an adaptive Funktionalitäten.

Die Ergebnisse der Anforderungsanalysen aus den Focus Groups und den Fragebogen ergaben, dass die verschiedenen Nutzergruppen auch unterschiedliche Anforderungen an ihr Handy stellten, was die Idee einer Personalisierung von Handys bestätigte. Auch der Prototyp wurde von den Teilnehmern der Focus Groups positiv bewertet. Bezüglich der Form der Personalisierung zeigte sich eine Tendenz hin zu einem adaptiven System. Die Adaptivität wurde aber bei vielen Teilnehmern an bestimmte Voraussetzungen geknüpft, wie z.B. die Bestimmung des Ausmaßes der Systemautonomie oder die Festlegung des Einsatzbereichs der Adaptivität.

#### **2.2.5 Forschungsprojekte der Hochschule der Medien Stuttgart und der Siemens AG CT IC7 zum Themenfeld Personalisierung und Adaptivität**

In einer Studienkooperation widmen sich die Hochschule der Medien in Stuttgart und die Siemens AG CT IC7 dem Forschungsfeld Personalisierung und Adaptivität. Ziel ist es sowohl allgemeingültige Erkenntnisse zu den Auswirkungen personalisierbarer und adaptiver Systeme zu gewinnen, als auch Anforderungen und Richtlinien für bestimmte Produkte und Produktklassen zu spezifizieren.

Im Sommersemester 2003 beschäftigte sich ein Seminar des Master-Studiengangs Informationswirtschaft der Hochschule der Medien in Stuttgart mit der Adaptivität von Interaktionselementen bei Microsoft Windows XP und Microsoft Word XP (Burmester, 2003). Die Studie beschäftigte sich mit folgenden Fragestellungen:

- Werden Expertennutzer mit umfassenden Nutzungserfahrungen von Microsoft XP Produkten durch adaptive Funktionalität eher behindert und Novizennutzer mit geringen Nutzungserfahrungen von Microsoft XP Produkten eher unterstützt?
- Verursacht Adaptivität Usability-Probleme und behindert gleichzeitig automatisierte Handlungsabläufe der Nutzer?

- Wie könnte Adaptivität gestaltet sein um die Effektivität und Effizienz der Softwarenutzung zu erhöhen?

Außerdem sollte durch den explorativen Studiencharakter eine Ergebnisbasis erarbeitet werden, die Ansätze für die Ableitung weiterer Untersuchungen ermöglicht. Die Studie fand als Usability-Test statt und adressierte den Erstkontakt einer adaptiven Nutzung entsprechend den von Mackay beschriebenen Einstiegshürden, die dadurch entstehen, dass Nutzer einen Grund benötigen, um ihr Produkt überhaupt personalisieren zu wollen (Mackay, 1991).

Zur Beantwortung der ersten Studienfrage erfolgte die Rekrutierung der Studienteilnehmer aufgrund ihrer Nutzungserfahrungen mit Microsoft XP Produkten. Die Teilnehmer wurden dementsprechend in die Gruppen Experten- und Novizennutzer unterteilt. Die Hälfte jeder Nutzergruppe arbeitete jeweils mit aktivierter Adaptivität, die andere Hälfte mit nicht-adaptiver Testsoftware. Um adaptive Auswirkungen exakt beurteilen zu können, wurde jeder Bedienungsschritt der Aufgaben genau vorgegeben. Der Testverlauf wurde protokolliert und durch Videomitschnitte sowie Aufzeichnungen eines Eye-Trackers ergänzt. Aus den Aufzeichnungen wurden objektive quantitative Daten zu Bearbeitungsschritten und -zeiten gewonnen. Die Probanden erhielten nach den Aufgabenblöcken jeweils einen Fragebogen, abschließend wurden sie bezüglich ihrer Eindrücke zur Aufgabenbearbeitung interviewt.

Betreffend der ersten Fragestellung konnte kein signifikanter Unterschied zu Auswirkungen der Nutzung adaptiver Systeme zwischen den beiden Nutzergruppen, Experten und Novizen, festgestellt werden. Die Ergebnisse zeigten über beide Nutzergruppen hinweg eine Tendenz zu einer subjektiv positiven Bewertung der Adaptivität, während die objektiven quantitativen Daten eine Tendenz zu einer höheren Effektivität und Effizienz des nicht-adaptiven Systems zeigten. Bei dem adaptiven System wurden sowohl längere Aufgabenbearbeitungszeiten als auch eine höhere Zahl an Usability-Probleme gemessen. Bezüglich auftretender Usability-Probleme wurde bei der adaptiven Gruppe jedoch beobachtet, dass ein Drittel aller Usability-Probleme in den ersten 5 von 34 Aufgaben auftraten, also ein Lerneffekt bezüglich der Funktionalität von Adaptivität bei den Nutzern einsetzte.

Zeitgleich zu der vorliegenden Ausarbeitung beschäftigten sich Studenten des Studiengangs Informationsdesign der Hochschule der Medien in Stuttgart mit gestalterischen Möglichkeiten zur Senkung von Einstiegshürden für die Nutzung adaptiver Systeme (Burmester, Wissmann, 2004). Die Seminarteilnehmer erarbeiteten Verbesserungsvorschläge, visuelle Szenarien und Prototypen für Interaktionselemente bei Microsoft Windows XP und Microsoft Office XP,

Informationssysteme in Automobilen, Mobiltelefone und Backöfen. Wesentliche Verbesserungspotentiale für adaptive Systeme sah man in einer besseren Erklärung von Vorteilen adaptiver Funktionen, einer besseren Darstellung von adaptiven Anpassungen und einer besseren Integration und Schnittstellengestaltungen von adaptiven und adaptierbaren Funktionen. Die Ergebnisse dienten für die Erarbeitung von Usability-Richtlinien zur Gestaltung von adaptiven und personalisierbaren UIs in einem abschließenden Workshop. Demnach sollten folgende Kriterien bei der Gestaltung berücksichtigt werden:

- Simplicity beschreibt die Einfachheit der Nutzung,
- Konsistenz beschreibt die innere Stimmigkeit des Systems,
- Kontrolle beschreibt die Kontrollierbarkeit des Systems,
- Attraktivität beschreibt den Spaß an der Nutzung,
- Nutzen beschreibt den Effekt auf das Ergebnis der Aufgabe,
- Nutzungskontext beschreibt die Angemessenheit von Anpassungen an die Aufgabenstellung, basierend auf der Analyse der Nutzergruppe des Nutzers, der Nutzungssituation, etc...,
- Integration beschreibt die Einbindung der adaptiven oder personalisierbaren Funktionalität in das System,
- Einfacher Zugriff beschreibt die Möglichkeit auf die adaptive oder personalisierbare Funktionalität zugreifen zu können,
- Vertrauen beschreibt das Vertrauen des Nutzers in das System,
- Unaufdringlichkeit / Etikette beschreibt die Darstellung und Veranschaulichung von Anpassungen und Informationen über die Anpassungen,
- Transparenz / Selbstbeschreibungsfähigkeit beschreibt die Möglichkeit für den Nutzer die Funktionsweise des Systems zu verstehen, vor allem im Bezug auf adaptive und personalisierbare Funktionen,
- Einstieg erleichtern / Hürden vermindern beschreibt die Voreingenommenheit gegenüber adaptiven und personalisierbaren Systemen zu reduzieren und die Nutzer durch Informationen über die Vorteile adaptiver und personalisierbarer Systeme zu ermuntern, von den Systemen Gebrauch zu machen.



### **2.3    *Problemstellung***

Wie in Kapitel 2.2.4 und 2.2.5 dargestellt, finden sich nur wenige Studien, die sich mit dem Forschungsfeld Personalisierung und Adaptivität auseinandersetzen.

In der theoretischen Diskussion spaltet sich die Forschungsgemeinde bezüglich der einzusetzenden Form zur Personalisierung eines Systems in zwei Lager, das erste argumentiert für die Adaptierbarkeit, das zweite bevorzugt die Adaptivität (Horvitz, 1999). Die bisherigen Forschungsergebnisse weisen auch nur ansatzweise Tendenzen in Bezug auf die Usability und die Nutzerpräferenzen aus (vgl. Kapitel 2.2.4 und 2.2.5). Außerdem basieren viele Studienergebnisse auf Labortests mit vordefinierten Aufgaben, was nur eingeschränkt die tatsächlichen Arbeitsgewohnheiten und Anforderungen der Nutzer widerspiegelt.

Für weiterführende Untersuchungen besteht die Notwendigkeit adaptive Systeme in Langzeitstudien zu evaluieren (Höök, 1997), da die Adaptivität erst nach einer gewissen Zeit der Nutzung den Umfang ihrer Wirksamkeit entfaltet. Für eine Langzeitevaluation empfiehlt es sich eine Feldstudie anstatt einer Laborstudie durchzuführen, da die Bearbeitung normalen Aufgaben in gewohnter Arbeitsumgebung die Wahrscheinlichkeit erhöht, dass Nutzer entsprechend ihrer tatsächlichen Bedürfnisse und Anforderungen personalisieren (McGrenere, Baecker, Booth, 2002). Feldstudien bieten für die Forschungsarbeit den Vorteil eine große Datenmenge generieren zu können ohne relevante Daten, die nicht von Beginn an im direkten Fokus der Untersuchung standen, zu verlieren. Allerdings ist der Vergleich von individuellen Aufgaben in unterschiedlichen Arbeitsumgebungen auch mit einem hohen Zeit-, Arbeits- und Kostenaufwand verbunden. Außerdem kann sich eine große Menge der Daten als gegebenenfalls nicht relevant bzw. die Suche nach relevanten Daten sich als schwierig erweisen (Millen, 2000).

Die größte Herausforderung bei der Untersuchung adaptiver Systeme stellt wohl die Abgrenzung der Auswirkungen einer adaptiven Funktionalität von der Gesamtgestaltung des Systems dar. Die meisten Studien vergleichen deshalb eine adaptive Version des Testprodukts mit einer Version ohne adaptive Funktionalität. Allerdings besteht das Problem, dass eine nicht-adaptive Systemvariante möglicherweise nicht optimal umgesetzt wurde, da das Konzept des UI auf eine vollständige Integration der Adaptivität abzielte. Das nicht-adaptive Systeme würde dem Nutzer somit nicht die Möglichkeit bieten eine Interaktion oder eine Aufgabe optimal

durchführen zu können. Somit kann auch dieser Vergleich nicht unbedingt eindeutig die Auswirkungen der Adaptivität abgrenzen (Höök, 1997).

Außerdem macht die gewollte Eigenschaft personalisierbarer Systeme auf den Nutzer angepasste Arbeitsinstrumente bereitzustellen, eine generelle Voraussage bezüglich der Entwicklung eines adaptiven Systems schwierig. Die sich so ergebenden spezifischen Anforderungen an adaptive Systeme können durch den Einsatz gängiger Methoden nur eingeschränkt evaluiert werden und erfordern das Design eigener, auf die Studienfragen zugeschnittener Erhebungsinstrumente, wie z.B. spezieller Fragebogen (Burmester, 2003), welche jedoch nicht auf ihre Reliabilität und Validität überprüft sind.

## **2.4 Studienfragen**

### **2.4.1 Studienziel**

In der vorliegenden Arbeit soll ein Konzept zur Untersuchung adaptiver UI-Elemente mit Methoden des UCD entwickelt werden. Ergebnis des Konzepts sind spezifische Erhebungsinstrumente zur Evaluation von adaptiven Systemen.

Die Studie untersucht die Auswirkungen von Adaptivität bei Interaktionselementen über einen längeren konstanten Gebrauch eines Produkts in der normalen Arbeitsumgebung der Nutzer. Entsprechend der in Kapitel 2.3 beschriebenen Problemstellung, sollen durch eine Aufteilung der Probanden in eine Adaptive Gruppe, die mit einem adaptiven System arbeitet, und einer Nicht-Adaptive Gruppe, die mit einem identischen nicht-adaptiven System arbeitet, Unterschiede zwischen adaptiven und nicht-adaptiven System herausgearbeitet werden.

### **2.4.2 Forschungsfragen**

Im Rahmen der Forschungsarbeit zum Thema Adaptivität der Hochschule der Medien und der Siemens AG CT IC7 kommt der Studie zur Adaptivität bei Interaktionselementen (Burmester, 2003) als Vorläufer der vorliegenden Untersuchung eine besondere Bedeutung bei der Bildung der Forschungsfragen zu. Aus den Ergebnissen, der in den Kapiteln 2.2.4 und 2.2.5 beschriebenen Studien zum Thema Adaptivität und dem in Kapitel 2.4.1 aufgestellten Studienziel, leiten sich für die Untersuchung die folgenden relevanten Fragestellungen ab.

### **Erste Forschungsfrage**

Bezüglich möglicher, durch die Adaptivität auftretender Usability-Probleme verfolgt die Untersuchung in einem ersten Schritt das Ziel, diese entsprechend ihrer Art, Häufigkeit und Auswirkungen zu identifizieren und mögliche Erklärungen für das Auftreten zu lokalisieren. Der Verlauf auftretender Usability-Probleme soll skizziert und eine Verbindung zwischen identifiziertem Problem, Zeitpunkt des Auftretens und Nutzungszeitraum hergestellt werden. Es soll untersucht werden, ob sich bei Adaptivität ein längerer Nutzungszeitraum im Bezug auf den Verlauf des Auftretens, der Häufigkeit und der Auswirkungen von Usability-Problemen entsprechend den Ergebnissen von Burmester (Burmester, 2003) positiv auswirkt.

### **Zweite Forschungsfrage**

Die Untersuchung soll klären, ob adaptive UI-Elemente sich für die Nutzung eines Systems entsprechend den Ergebnissen einer besseren Erlernbarkeit von personalisierbaren gegenüber nicht-personalisierbaren Systemen von McGrenere, Baecker und Booth (McGrenere, Baecker, Booth, 2002) lernfördernd auswirken.

### **Dritte Forschungsfrage**

Die dritte Fragestellung greift auf die Ergebnisse von Burmester (Burmester, 2003) zurück, dass die Probanden das adaptive System subjektiv positiver bewerten als das System ohne Adaptivität und sich Adaptivität wünschen. Die vorliegende Studie soll diese Ergebnisse untersuchen.

### **Vierte Forschungsfrage**

Basierend auf der dritten Fragestellung und den geschilderten Ergebnissen von Burmester (Burmester, 2003) soll die Frage geklärt werden, ob adaptive Funktionen sich positiv auf die Wahrnehmung von pragmatischer und hedonischer Qualität sowie der daraus resultierenden Attraktivitätsbewertung von Seiten der Nutzer auswirken. Dabei bezeichnet die pragmatische Qualität das Erreichen eines Nutzungsziels. Hedonische Qualität beschreibt eine nicht aufgabenbezogene Qualität bezüglich der Aktualität eines Produkts, der Identifikation mit einem Produkt und dem Wunsch nach Kommunikation und Selbstdarstellung durch das

Produkt. Die globale Bewertung findet sich in der Wahrnehmung der Attraktivität eines Produkts (Hassenzahl, Burmester, Koller, 2003).

### **Fünfte Forschungsfrage**

Die Studie zur Adaptivität bei Interaktionselementen von Burmester (Burmester, 2003) beschäftigte sich mit Unterschieden in der Behinderung und Unterstützung durch adaptive Funktionalität zwischen Nutzern mit unterschiedlichen Nutzungserfahrungen. Es wurde die These aufgestellt, dass sich Adaptivität wahrscheinlich hindernd auf die Arbeit von Nutzern mit hoher Nutzungserfahrung und unterstützend auf die Arbeit von Nutzern mit geringer Nutzungserfahrung auswirkt. Es konnte jedoch kein signifikanter Unterschied zwischen beiden Nutzergruppen festgestellt werden. Diese Ergebnisse sollen in der vorliegenden Arbeit überprüft werden.

#### **2.4.3 Studientyp und Gültigkeitsanspruch**

Die Studie hat einen weitgehend deskriptiv-explorativen Charakter und soll der Verbesserung von Verfahren zur Evaluation von adaptiven Systemen dienen. Trotz des explorativen Vorgehens bei der Erstellung des Konzepts, hat die Arbeit auch zum Ziel Ergebnisse anderer Studien zu überprüfen und weiterführende Erkenntnisse zu den aufgestellten Forschungsfragen zu liefern (vgl. Kapitel 2.4.2). Ergebnisse und Erkenntnisse sollen in ihrer Form möglichst allgemeingültig für die Adaptivität von Interaktionselementen und übertragbar auf Produkte und softwarebasierte Dienste sein, die durch adaptive Interaktionselemente erweitert werden könnten, wie beispielsweise Handys, Haushaltsgeräte oder Informationssysteme.

#### **2.4.4 Variablen der Studie**

##### **Unabhängige Variablen (Einflussgrößen)**

Für die vorliegende Studie wurden, bezogen auf die Probanden, folgende unabhängige Variablen festgelegt:

- Demographische Daten: Alter, Geschlecht, Ausbildung und Beruf,
- Zuordnung in eine Nutzergruppe aufgrund der Nutzungserfahrungen,
- Nutzungshäufigkeit des Testprodukts,

- Tätigkeitsbeschreibung und Dauer einer Aufgabe.

Im Kontext der Software bilden relevante unabhängige Variablen die Aktivierung der Adaptivität und das Setup der adaptiven Funktionalität.

### **Abhängige Variablen (Zielgrößen)**

Abhängige Variablen, im Kontext der vorliegenden Studie, repräsentieren vor allem die subjektiven Empfindungen und Einschätzung der Testperson bezüglich des Arbeitens mit der Testsoftware. Die Schwierigkeit der Datenanalyse lag in möglichen Abweichungen zwischen subjektiven Angaben der Testperson und beobachteten Einschätzungen des Testleiters. Wichtige Variablen sind:

- Einstellung gegenüber Formen der Personalisierung,
- Inhaltliche Schwierigkeit einer Aufgabe und erlebte mentale Anstrengung bei der Aufgabenbearbeitung,
- Usability-Probleme und unterstützende Effekte,
- Anforderungen an adaptive Systeme,
- Lernförderlichkeit,
- Effizienz,
- Attraktivität und hedonische Qualitäten.

### **Andere Variablen (Störgrößen)**

Als Störgrößen kommen äußere Einflüsse während der Bearbeitung einer Aufgabe in Betracht, die sich nicht auf die adaptive Funktionalität zurückführen lassen, wie etwa die Unterbrechung durch einen Telefonanruf oder Probleme mit dem Computersystem. Sie beeinflussen jedoch in bestimmten Situationen entscheidend den Verlauf der Aufgabenbearbeitung, sollten also zur möglichen Erklärungsfindung bei individuellen Nutzungssituationen berücksichtigt werden.

## 3 Methodik

### 3.1 Untersuchungsart

Die Untersuchung wurde als Feldstudie durchgeführt mit dem Ziel Auswirkungen der Adaptivität von Interaktionselementen über einen längeren Zeitraum konstanter Nutzung zu untersuchen. Die Probanden arbeiteten in ihrem gewohnten Arbeitsumfeld und erledigten ihre normalen Aufgaben. Als Evaluationsverfahren zum Einsatz kam die Methode der Tagebuchstudie, ergänzt durch periodische Interventionen vor Ort bei den Probanden. Die Studie endete mit einer Abschlussbefragung.

Da die Forschungsfragen vor allem an die subjektiven Einschätzungen und Erfahrungen von Nutzern mit adaptiven Systemen adressiert waren, sah das Studienkonzept auch hauptsächlich Instrumente zur Erhebung subjektiver Daten vor.

### 3.2 Erhebungsinstrumente

#### 3.2.1 Einstellung gegenüber Formen der Personalisierung

Die Einstellung gegenüber Formen der Personalisierung beschreibt den Wunsch der Nutzer nach einem System, das auf ihre Arbeitsgewohnheiten angepasst ist. Für die Datenerhebung wurden den Nutzern mehrere Antwortmöglichkeiten vorgegeben, Mehrfachauswahl war möglich, ein Hinweis erfolgte durch einen Zusatz zur Fragestellung, wie Abbildung 2 illustriert:

***Würden Sie sich wünschen, dass ... Ihren individuellen Arbeitsgewohnheiten angepasst sind?***

Mehrfachauswahl möglich...

- ☐ Ja, sollten sich **automatisch** anpassen, ...
- ☐ ...auch **ohne mein Zutun**.
- ☐ ...aber nur **wenn ich informiert werde** über die vorgenommenen Anpassungen.
- ☐ ...aber nur mit meiner **ausdrücklichen Zustimmung**.
- ☐ Ja, durch die Möglichkeit, **selbst Anpassungen vornehmen** zu können.
- ☐ Nein, ich **bin zufrieden** mit der Standardversion.

**Abbildung 2 - Erhebungsinstrument Einstellung gegenüber Formen der Personalisierung**

Die erste Auswahlmöglichkeit entspricht dem Wunsch nach einem adaptiven System, eine Antwortspezifizierung bezüglich der Systemautonomie war notwendig und wie folgt gegliedert:

- Völlige Systemautonomie: Automatische Anpassungen erfolgen ohne Zutun des Nutzers.
- Weitgehende Systemautonomie: Die Anpassungen erfolgen automatisch, der Nutzer möchte aber über die Anpassungen informiert werden.
- Kontrollierte Systemautonomie: Die Anpassungen können nur mit der Bestätigung des Nutzers vom System durchgeführt werden.

Die zweite Antwort nennt die Möglichkeit, das System entsprechend eigener Vorstellungen, selbst anpassen, adaptieren zu können.

Der Wunsch nach einer Standardversion schließt durch die Formulierung als Verneinung, Kombinationen mit den anderen Antwortvorgaben und somit die Möglichkeit zur Personalisierung explizit aus.

Zum Einsatz kam das Erhebungsinstrument im Fragebogen zum Studienbeginn (vgl. Kapitel 3.3.1) und im Fragebogen zum Studienabschluss (vgl. Kapitel 3.3.5.1).

### **3.2.2 Skala zur Beurteilung subjektiver mentaler Anstrengung (BSMA) zur Messung der mentalen Anstrengung einer Aufgabenbearbeitung**

Die mentale Anstrengung beschreibt die Energieleistung, die ein Nutzer bei der Aufgabenbearbeitung einsetzen muss, um sein Ziel zu erreichen (Arnold, 1999). Das Messen der mentalen Anstrengung hat somit den Vorteil, dass sich die Ergebnisse auf die unmittelbar bearbeitete Aufgabe beziehen. Zur Erfassung der mentalen Anstrengung wurde die Skala zur Beurteilung subjektiver mentaler Anstrengung (BSMA) (Eilers, Nachreiner, Hänecke, 1986) eingesetzt, die deutsche Version des Subjective Mental Effort Questionnaire (SMEQ).

Allerdings vermengen sich bei der Messung mit dem SMEQ oder dem BSMA unter Umständen die Gebrauchstauglichkeit des Produkts und die inhaltliche Aufgabenschwierigkeit (Hassenzahl, 2004). Bei einer Laborstudie ist dieser Aspekt durch das Vorgeben von Aufgaben relativierbar, diese Möglichkeit besteht bei einer Feldstudie jedoch nicht. Damit bei der Untersuchung die inhaltlichen Aspekte der Aufgabe, die nicht mit der Gebrauchstauglichkeit der Software zusammenhängen, aus der Messung der mentalen Anstrengung herausgerechnet werden

konnten, wurde zusätzlich zum BSMA ein Erhebungsinstrument zur Messung der Aufgabenschwierigkeit erstellt. Wie Abbildung 3 zeigt, illustrierte eine einleitende Erklärung mit einem Beispiel die Fragestellung, die Messung erfolgte durch eine 7-stufige bipolare Skala mit dem gegensätzlichen Wortpaar „Einfach - Kompliziert“ als extreme Ausprägungen.

**Bitte beurteilen Sie die Schwierigkeit Ihrer Aufgabe:**

Mit der Schwierigkeit Ihrer Aufgabe ist hier gemeint, wie einfach oder kompliziert Sie z.B. das Anfertigen einer englischen Übersetzung einschätzen aufgrund Ihrer englischen Sprachkenntnisse. **Nicht** gemeint ist hingegen, wie schwierig Ihnen die Erledigung dieser Aufgabe mit ... gefallen ist!

Einfach ☐ ☐ ☐ ☐ ☐ ☐ ☐ Kompliziert

**Abbildung 3 - Erhebungsinstrument Aufgabenschwierigkeit**

Beide Erhebungsinstrumente, BSMA und Aufgabenschwierigkeit, wurden zur direkten Bewertung einer Aufgabenbearbeitung im Tagebuch (vgl. Kapitel 3.3.2) eingesetzt. In den Fragebogen zur Intervention (vgl. Kapitel 3.3.3.2) und zum Studienabschluss (vgl. Kapitel 3.3.5.1) erfolgte die Bewertung als Mittel über einen Nutzungsabschnitt. Dies ermöglichte eine anschaulichere Skizzierung des Bewertungsverlaufs. Bei Anwendung in den Fragebogen wurde die Formulierung der Erhebungsinstrumente auch um entsprechende Erläuterungen ergänzt, dass es sich um eine Bewertung im Mittel handelte.

### **3.2.3 Usability-Probleme und unterstützende Effekte**

Adaptivität wirkt sich durch die Anpassungen des Interaktionselements stark auf die Arbeit des Nutzers mit der Software aus. Die erste Fragestellung (vgl. Kapitel 2.4.2) sollte klären, ob sich dies eher hindernd in Form von Usability-Problemen oder unterstützend auf die Arbeit des Nutzers auswirkt. Da das Studienkonzept nicht auf die Erhebung von objektiven quantitativen Daten zur Beantwortung der Forschungsfragen abzielte, musste ein Weg gefunden werden auftretende Usability-Probleme und unterstützende Effekte in die, auf subjektiven Einschätzungen der Probanden basierende, Datenerhebung einzubeziehen.

Vor der Studie wurden anhand einer Voruntersuchung der Testsoftware mögliche interaktionsadaptive Effekte während der Nutzung des Interaktionselements dokumentiert und als „Ereignisse“ definiert. Wie in Abbildung 4 zeigt, wurden diese „Ereignisse“ in der Fragestellung der Erhebungsinstrumente möglichst präzise beschrieben und anhand ihrer Wirkung auf den Nutzer in die beiden Ereigniskategorien „Unterstützende Ereignisse“ und „Behindernde Ereignisse“ aufgeteilt:



### Unterstützende Ereignisse

Ich konnte den gesuchten **Menüeintrag** schnell dort finden, **wo** ich ihn **erwartet** hatte.

Ich konnte den Menüeintrag schnell finden, da das **Menü kurz und übersichtlich** war.

Ich konnte den **Menüeintrag** schnell finden, da er **prominent platziert** war.

### Behindernde Ereignisse

Ich konnte den **Menüeintrag nicht** dort finden, **wo** ich ihn **erwartet** hatte.

Ich musste den Menüeintrag erst suchen, da das **Menü lang und unübersichtlich** war.

Ich konnte den **Menüeintrag gar nicht finden**. Stattdessen bin ich auf ein anderes Interaktionselement (Symbole, Shortcuts, Assistent) ausgewichen.

Um den Menüeintrag zu finden, musste ich das **Menü verlängern**.

## Abbildung 4 - Ereignisse

Aufgrund ihrer Auswirkungen können die „Behindernden Ereignisse“ als „Usability-Probleme“, die „Unterstützenden Ereignisse“ als „Unterstützende Effekte“ bei der Nutzung des Interaktionselements bezeichnet werden.

In einem weiteren Schritt wurden relevante Parameter bezüglich der definierten „Ereignisse“ festgelegt:

- Die Anzahl eines „Ereignisses“ beschreibt die absolute Häufigkeit mit der „Ereignisse“ und die Ereigniskategorien durchschnittlich von den Probanden während der Aufgabenbearbeitung wahrgenommen wurden. Zur Erhebung sollten die Probanden ein aufgetretenes „Ereignisses“ durch das Ankreuzen eines dem „Ereignis“ zugeordneten Kästchens markieren. Das Ergebnis der Auszählung markierter „Ereignisse“ wurde durch die Aufgabenzahl z.B. eines Nutzungsabschnitts dividiert, so ergab sich die Wahrscheinlichkeit, mit der ein „Ereignis“ oder eine Ereigniskategorie pro Aufgabe in einem Nutzungsabschnitt auftrat.
- Die geschätzte Häufigkeit einer Ereigniskategorie je Aufgabe beschreibt in welchem Verhältnis zu den Gesamtinteraktionen eine Ereigniskategorie während der Aufgabenbearbeitung von den Nutzern wahrgenommen wurde. Die Erhebung fand im direkten Bezug zu jeder Ereigniskategorie mittels der einleitenden Frage „Wie oft sind diese Ereignisse bei der Bearbeitung Ihrer Aufgabe aufgetreten?“ und der 5-stufigen Ratingskala zur Häufigkeit von Rohrmann (Bortz, Döring, 2003, S.178-179) statt.
- Durch die Intensität einer Ereigniskategorie schätzten die Studienteilnehmer wie stark sich eine Ereigniskategorie auf ihre Arbeit mit dem Interaktionselement auswirkte. Für die „Usability-Probleme“ beschreibt die Intensität somit die Schwere der Behinderung

und für die „Unterstützenden Effekte“ den Grad der Unterstützung. Die Erhebung fand im direkten Bezug zu jeder Ereigniskategorie mittels der einleitenden Frage „Haben diese Ereignisse Sie bei der Bearbeitung Ihrer Aufgabe unterstützt bzw. behindert?“ und der 5-stufigen Ratingskala zur Intensität von Rohrmann (Bortz, Döring, 2003, S.178-179) statt.

Die Datenerhebung geschah im unmittelbaren Bezug zur Aufgabenbearbeitung im Tagebuch (vgl. Kapitel 3.3.2).

### **3.2.4 Anforderungen an adaptive Systeme**

Entsprechend der in Kapitel 2.2.3 erläuterten Anforderungen an adaptive Systeme wurden die folgenden Dimensionen zur Untersuchung adaptiver Auswirkung der Systeminteraktion auf die Nutzung erstellt:

- Nutzen beschreibt den Effekt auf das Ergebnis der Aufgabe durch die Nutzung des Interaktionselements,
- Etiquette beschreibt die Selbstbeschreibungsfähigkeit sowie die Darstellung und Veranschaulichung des Interaktionselements,
- Privatsphäre beschreibt den Eingriff des Systems in die Privatsphäre des Nutzers durch die Interaktionen,
- Verständnis beschreibt das Verständnis des Interaktionselements durch den Nutzer,
- Vertrauen beschreibt das Vertrauen des Nutzers in das System aufgrund der Interaktionen,
- Breadth of Experience beschreibt die Erfüllung von Bedürfnissen eines Nutzers durch die Gegebenheiten des Interaktionselements,
- Kontrolle beschreibt die Kontrollierbarkeit des Interaktionselements.

Ein Begriffspaar extremer Gegenpole, ergänzt durch einen einleitenden Text, bildete bei der Datenerhebung die jeweilige Dimension in einer 5-stufige Skalen ab.

**Nutzen**

Die Ereignisse haben mich bei der Erledigung meiner Aufgabe...

Unterstützt ☐ ☐ ☐ ☐ ☐ Behindert**Etiquette**

Die Darstellung / Veranschaulichung der Ereignisse empfand ich als...

Angenehm ☐ ☐ ☐ ☐ ☐ Störend**Privatsphäre**

Aufgrund der Ereignisse sehe ich meine Privatsphäre durch das System...

Gewahrt ☐ ☐ ☐ ☐ ☐ Gefährdet**Verständnis**

Generell ist das Auftreten von Ereignissen...

Vorhersagbar ☐ ☐ ☐ ☐ ☐ Unberechenbar**Vertrauen**

Das Systemverhalten empfinde ich als...

Zuverlässig ☐ ☐ ☐ ☐ ☐ Unzuverlässig**Breadth of Experience**

Meine Bedürfnisse werden vom System...

Verstanden ☐ ☐ ☐ ☐ ☐ Nicht verstanden**Kontrolle**

Ich fühle mich dem System gegenüber...

In Kontrolle ☐ ☐ ☐ ☐ ☐ Ausgeliefert**Abbildung 5 - Erhebungsinstrument Anforderungen an adaptive Systeme**

Die Abfrage, entsprechend der Darstellung in Abbildung 5, erfolgte immer im Bezug zu aufgetretenen „Ereignissen“ (vgl. Kapitel 3.2.3). Die Abfrage fand jeweils zur direkten Bewertung einer Aufgabenbearbeitung im Tagebuch (vgl. Kapitel 3.3.2) und zum Zweck der Veranschaulichung des Bewertungsverlaufs in den Fragebogen zur Intervention (vgl. Kapitel 3.3.3.2) und zum Studienabschluss (vgl. Kapitel 3.3.5.1) als Mittel über einen Nutzungsabschnitt statt. Ein einleitender Text stellte den Zusammenhang mit den „Ereignissen“ her, Erläuterungen zur Bewertung im Mittel ergänzten die Formulierungen in den Fragebogen. Der Erhebung der Dimension Nutzen erfolgte in der beschriebenen Form jedoch nur in den Fragebogen. Um einen Vergleichswert für die Bewertung aus dem Tagebuch zu erhalten, wurden die Werte der Intensität von „Unterstützenden und Behindernden Ereignissen“ (vgl. Kapitel 3.2.3) im gleichen Verhältnis miteinander verrechnet.

Im Fragebogen zum Studienabschluss (vgl. Kapitel 3.3.5.1) wurden die Dimensionen der Anforderungen an adaptive Systeme sowohl in der beschriebenen Form, im Bezug zu den „Ereignissen“ als auch in einer überarbeiteten Version zur finalen Einschätzung und für mögliche Verbesserungspotentiale abgefragt. Die Dimensionen der zweiten Fragestellung wurden ergänzt durch die Usability-Richtlinien für adaptive und personalisierbare UIs (vgl. Kapitel 2.2.5) der Seminargruppe zum Thema Gestaltungsmöglichkeiten von adaptiven

Systemen der Hochschule der Medien in Stuttgart (Burmester, Wissmann, 2004). Die erweiterte Abfrage beinhaltete folgende Dimensionen:

- Simplicity beschreibt die Einfachheit der Nutzung des Interaktionselements,
- Konsistenz beschreibt die innere Stimmigkeit des Interaktionselements,
- Kontrolle beschreibt die Kontrollierbarkeit des Interaktionselements,
- Attraktivität beschreibt den Spaß an der Nutzung des Interaktionselements,
- Nutzen beschreibt den Effekt auf das Ergebnis der Aufgabe durch die Nutzung des Interaktionselements,
- Nutzungskontext beschreibt die Angemessenheit des Interaktionselements an die Aufgabenstellung,
- Integration beschreibt die Einbindung der adaptiven Funktionalität in das System,
- Zugriff beschreibt die Möglichkeit auf die adaptive Funktionalität zugreifen zu können,
- Vertrauen beschreibt das Vertrauen des Nutzers in das System aufgrund der Interaktionen,
- Etikette beschreibt die Selbstbeschreibungsfähigkeit sowie die Darstellung und Veranschaulichung des Interaktionselements.

Alle Dimensionen wurden anhand einer einführenden Frage und der 5-stufigen Ratingskala zur Bewertung nach Rohrmann (Bortz, Döring, 2003, S.178-179) abgefragt. Wie in Abbildung 6 anhand der Dimension Simplicity veranschaulicht, bot das zusätzliche Feld „Keine Angabe“ die Möglichkeit die Frage als irrelevant einzustufen.

Die Handhabung ... sollte einfacher sein.

Völlig Falsch ☐ ☐ ☐ ☐ ☐ Völlig richtig Keine Angabe ☐

**Abbildung 6 - Erhebungsinstrument Anforderungen an adaptive Systeme Abschlussbefragung Beispiel Dimension Simplicity**

Eine vorgelagerte offene Frage zu allgemeinen Verbesserungsvorschlägen des Testprodukts sollte äußere Faktoren der allgemeinen Gestaltung und Gebrauchstauglichkeit von der eigentlichen Bewertung der Anforderungen an die Adaptivität trennen. Außerdem konnten weitere Verbesserungsvorschläge zum Interaktionselement in einer offenen Frage im Anschluss an die Abfrage der Dimensionen gemacht werden.

### 3.2.5 Lernförderlichkeit

Zur Messung der Lernförderlichkeit wurden Fragen zu Lernförderlichkeit und Erlernbarkeit der Standardfragebogen ISO-Norm 9241-10 (Prümper, 1997), IsoMetrics (Hamborg, Willumeit, Gediga, 1996) und CSUQ (Lewis, 1995) herangezogen und folgende Dimensionen, mit Bezug auf die Adaptivität, für das Erstellen eigener Fragen berücksichtigt:

- Zeit, die zum Erlernen einer Software benötigt wird,
- Anregung zur Exploration einer Software durch deren Gestaltung,
- Menge und Details, die erlernt werden müssen, um mit der Software entsprechend eigener Vorstellungen arbeiten zu können,
- Einpräbarkeit von Gelerntem einer Software.

Die Abfrage der Dimensionen erfolgte anhand eines einleitenden Satzes und der 5-stufigen Ratingskalen zur Bewertung nach Rohrmann (Bortz, Döring, 2003, S.178-179) entsprechend Abbildung 7:

#### **Zeit**

Erfordert viel Zeit zum Erlernen.

Völlig falsch ☐ ☐ ☐ ☐ ☐ Völlig richtig

#### **Exploration**

Ermutigt dazu, auch mal neue Funktionen auszuprobieren.

Völlig falsch ☐ ☐ ☐ ☐ ☐ Völlig richtig

#### **Menge und Details**

Erfordert, dass man sich viele Details merken muss.

Völlig falsch ☐ ☐ ☐ ☐ ☐ Völlig richtig

#### **Einpräbarkeit**

Ist so gestaltet, dass sich einmal Gelerntes gut einprägt.

Völlig falsch ☐ ☐ ☐ ☐ ☐ Völlig richtig

### **Abbildung 7 - Erhebungsinstrument Lernförderlichkeit**

Zum Einsatz kam die Abfrage in den Fragebogen zur Intervention (vgl. Kapitel 3.3.3.2) und zum Studienabschluss (vgl. Kapitel 3.3.5.1).

### 3.2.6 Effizienz

Die Effizienz wurde anhand einer einführenden Frage mittels der 5-stufigen Ratingskala zur Bewertung nach Rohrmann (Bortz, Döring, 2003 S.178-179) entsprechend Abbildung 8 abgefragt.

Mit ... kann ich meine Aufgaben effizient erledigen  
Völlig falsch ☐ ☐ ☐ ☐ ☐ Völlig richtig

#### Abbildung 8 - Erhebungsinstrument Effizienz

Die Abfrage fand in den Fragebogen zur Intervention (vgl. Kapitel 3.3.3.2) und zum Studienabschluss (vgl. Kapitel 3.3.5.1) statt.

### 3.2.7 AttrakDiff zur Messung der Attraktivität und der hedonischen Qualität

Für die Datenerhebung zur Messung der Attraktivität und hedonischer Qualität wurde der von Marc Hassenzahl entwickelte Fragebogen AttrakDiff verwendet (Hassenzahl, Burmester, Koller, 2003).

Die Abfrage fand in den Fragebogen zur Intervention (vgl. Kapitel 3.3.3.2) und zum Studienabschluss (vgl. Kapitel 3.3.5.1) statt.

### 3.2.8 Qualitative Daten

Primäres Ziel der Erhebung qualitativer Daten war, die Ergebnisse der erstellten Erhebungsinstrumente durch verbale Äußerungen der Probanden während oder nach der Aufgabenbearbeitung sowie durch objektive Beobachtungen zu ergänzen. Individuelle Nutzungsweisen, spezielle Merkmale des UI oder äußere Einflüsse dienten zur Analyse individuellen Nutzungsverhaltens sowie der Erklärungsfindung bei einzelnen Nutzungssituationen. Ein weiteres Ziel bestand in der Beurteilung der für die Studie erstellten Erhebungsinstrumente zur Evaluation adaptiver Systeme durch Feedback der Probanden zum Studienaufbau und -verlauf.

Die Erhebung qualitativer Daten erfolgte durch offene Fragen, Protokollierung, *Recall Surveys*, der Befragung von Probanden zu bereits vergangenen Vorkommnissen (Rondoni, 2003), und Interviews.

Offene Fragen fanden sich immer am Ende des Tagebuchs (vgl. Kapitel 3.3.2), der Fragebogen zur Intervention (vgl. Kapitel 3.3.3.2) und zum Studienabschluss (vgl. Kapitel 3.3.5.1). Die Fragestellung bot die Möglichkeit zu Äußerungen bezüglich der in den Erhebungsinstrumenten abgegebenen Bewertungen, weiteren Einschätzungen zu der Aufgabenbearbeitung, äußeren Einflüssen, Feedback zum Studienaufbau und -ablauf und allen weiteren Eindrücken der Probanden während der Aufgabenbearbeitung.

Durch die Protokollierung während der Vor-Ort-Beobachtungen (vgl. Kapitel 3.3.3.1) wurden qualitative Daten erhoben, die sich auf Beobachtungen zur Aufgabenbearbeitung der Nutzer bezogen. Dies betraf allgemeine Beobachtungen zum Nutzerverhalten, wie beispielsweise individuelle Nutzungsweisen, spezielle Merkmale des UI oder äußere Einflüsse. Eine detaillierte Erfassung erfolgte zu Beobachtungen und Reaktionen sowie verbalen Äußerungen der Probanden aufgrund von Interaktionen, speziell dem Auftreten von „Usability-Problemen“ und „Unterstützenden Effekten“. Diese Beobachtungen dienten dazu die Einschätzungen der Probanden im Studienverlauf zu evaluieren.

Qualitative Daten aus Recall Surveys (vgl. Kapitel 3.3.3.3) sollten Tagebucheinträge mit abweichenden Bewertungen und von den Probanden als untypisch bezeichnete Aufgabenbearbeitungen analysieren oder auf Fragen der Testpersonen eingehen.

Die Erhebung qualitativer Daten im Interview zum Studienabschluss (vgl. Kapitel 3.3.5.2) diente als Ergänzung des Fragebogens zum Studienabschluss (vgl. Kapitel 3.3.5.1). Die Probanden wurden gebeten vor allem ihre Arbeitsweise, ihre Wahrnehmung und Motivation zur Adaptivität und Personalisierung darzustellen. Ein Abgleich dieser Einschätzungen mit den Beobachtungsprotokollen und den Daten der Erhebungsinstrumente ermöglichte eine detailliertere Erklärungsfindung für individuelle Nutzungssituationen.

### **3.2.9 Objektive quantitative Daten**

Da das Studienkonzept auf der Erhebung subjektiver Einschätzungen basierte, diente die Erhebung objektiver quantitativer Daten nur zur Ergänzung der Ergebnisse aus den erstellten Erhebungsinstrumenten, zur Analyse individuellen Nutzungsverhaltens sowie der Erklärungsfindung bei einzelnen Nutzungssituationen.

Eine Erhebung fand als Dokumentation der Benutzungsoberfläche (vgl. Kapitel 3.3.4) sowie stichprobenartig in Form einer A/V-Aufzeichnung bei der Vor-Ort-Beobachtung (vgl. Kapitel 3.3.3.1) statt.

Die Dokumentation der Benutzungsoberfläche verglich das Aussehen des Interaktionselements über den Studienzeitraum von Nutzungsabschnitt zu Nutzungsabschnitt. So ließen sich die adaptiven Anpassungen während eines Nutzungsabschnitts erkennen und auszählen. Dies gab Aufschluss über den tatsächlichen Umfang adaptiver Auswirkungen, welche für individuelle Bewertungen eine große Rolle spielen.

Die bei der A/V-Aufzeichnung gewonnenen Daten bestanden aus den Arbeitsschritten und Bearbeitungszeiten einer Interaktion.

### **3.3     *Einsatz der Erhebungsinstrumente***

#### **3.3.1   Fragebogen zum Studienbeginn**

Der Fragebogen zum Studienbeginn (vgl. Anhang B) diente der Erhebung relevanter Daten für die Auswahl und Rekrutierung der Studienteilnehmer. Abfragt wurden Kontaktdaten, demographische Daten, Nutzungserfahrungen, die Einstellung gegenüber Formen der Adaptivität (vgl. Kapitel 3.2.1) und Fragen bezüglich der Nutzung des geplanten Testprodukts.

Vor dem eigentlichen Fragebogen hieß ein kurzer Einleitungstext die Studienbewerber willkommen, schilderte ihnen kurz das Ziel und den Termin der Studie sowie Voraussetzungen zur Studienteilnahme. Ein zweiter Textabschnitt erläuterte in wenigen Worten den geplanten Studienablauf.

Unter dem Punkt „Persönliche Angaben“ wurden Kontaktdaten und Teile der demographischen Daten abgefragt. Die Bewerber sollten ihren Namen, Wohnort, die Adresse des Standortes ihres PCs, als Ort an dem Studie durchgeführt werden sollte, Telefonnummer und E-Mail-Adresse angeben. Für die Demographie erfolgte die Bestimmung des Geschlechts aufgrund der Namensangabe. Angaben zum Alter wurden anhand von vier Altersgruppen abgefragt:

- Bis 29 Jahre,
- 30 bis 39 Jahre,
- 40 bis 54 Jahre,
- ab 55 Jahren.



Im nächsten Abschnitt wurde nach der Ausbildung in Form verschiedener Kategorien gefragt, teilweise anhand von Beispielen illustriert:

- Ausbildungsberuf / Volontariat,
- Weiterführende Ausbildung (Meister, Techniker, ...),
- Hochschulstudium,
- Weiterführende akademische Ausbildung (Promotion, PhD-Programm, Habilitation, ...),
- Andere / Sonstige.

Die Studienbewerber hatten außerdem noch die Möglichkeit ihre Angaben zu spezifizieren und ihre derzeitige Tätigkeit in einem zusätzlichen Feld anzugeben.

Die Ermittlung der Nutzungserfahrung basierte auf den Angaben zu PC-Kenntnissen und Erfahrungen mit den Produktfamilien der Testsoftware. Für die Angabe der PC-Kenntnisse war eine Mehrfachauswahl der Kategorien

- Betriebssysteme,
- Datenbanken,
- Internet,
- Netzwerke,
- Office-Anwendungen,
- Programmierung,
- Sonstige Anwendungen

möglich, sowie eine Spezifizierung der Kenntnisse in einem Textfeld. Die Angabe bezüglich Erfahrungen mit den Produktfamilien gliederte sich in die vier aktuellsten Versionen sowie der Zusammenfassung älterer Versionen der Testsoftware und fragte nach der derzeitigen Nutzung, geschäftlich und privat.

Weitere Fragen nach der Nutzung des geplanten Testprodukts betrafen die Nutzungshäufigkeit, die Tätigkeitsbeschreibung und die Verwendung von Interaktionselementen. Die Nutzungshäufigkeit konnte angegeben werden als:

- Täglich, über 4 Stunden,
- Täglich, aber unter 4 Stunden,
- Mehrmals wöchentlich,
- Nicht mehr als einmal pro Woche.

Die Formulierung der Frage nach der Tätigkeitsbeschreibung erfolgte nach der Art ausgeführter Aufgaben mit dem geplanten Testprodukt und wurde gemessen mittels einer 5-stufigen Ratingskala mit dem gegensätzlichen Wortpaar „Meist gleiche Aufgaben - Meist wechselnde Aufgaben“. Bei der Frage nach der Verwendung von Interaktionselementen wurden nur Interaktionselemente berücksichtigt, die den vollen Funktionsumfang der Testsoftware darstellen können. Andere Interaktionselemente, die kontextsensitiv Ausschnitte des Funktionsumfangs abbilden, fanden keine Berücksichtigung. Die Liste verwendeter Interaktionselemente bestand somit aus folgenden Interaktionselementen:

- Menüs,
- Symbole,
- Tastatur-Shortcuts,
- Assistent.

Zur besseren Veranschaulichung illustrierten Grafiken und ggf. ein Beispiel die betroffenen Interaktionselemente.

Die letzten Seiten des Fragebogens beinhalteten eine Einverständniserklärung zum gegenseitigen Datenschutz, Kontaktdaten und eine Danksagung.

### **3.3.2 Feldstudie (Tagebuch)**

Als Evaluationsverfahren der Feldstudie kam die Tagebuchstudie zum Einsatz. Die Entscheidung hierzu fiel aufgrund des Alltagsbezugs der erhobenen Daten, der sehr kontextorientierten Einsatzmöglichkeit, der Generierung von Daten zu subjektiven Einschätzungen und Erfahrungen sowie der Möglichkeit, Verlaufskurven zu zeitlichen Veränderungen und individuellen Unterschieden gut zeichnen zu können (Bolger, Davis, Rafaeli, 2002).

Die Teilnehmer wurden angehalten bei einer definierten Begebenheit einen Eintrag in das Tagebuch vorzunehmen (Colbert, 2001). Diese Begebenheit war definiert als die Erledigung einer Arbeitsaufgabe mit der Testsoftware. Der Tagebucheintrag sollte unmittelbar nach Beendigung der Aufgabe vorgenommen werden. Diese Festlegung bot einen objektiven Vergleichsrahmen für die Analyse der Nutzung, hatte allerdings den Nachteil, die mögliche Fülle an Interaktionen und adaptiven Anpassungen nicht im Detail abbilden zu können oder in manchen Fällen eventuell gar keine relevanten Daten zu erheben. Um die Abbildung von Interaktionen und adaptiven Anpassungen pro Eintrag darstellbar zu halten, wurde ein zeitlicher

Rahmen von einer Stunde pro Aufgabe festgelegt. Dauerte eine Aufgabebearbeitung länger als eine Stunde sollte für die Folgezeit ein weiterer Eintrag in das Tagebuch vorgenommen werden, nach einer weiteren Stunde ein weiterer, usw. ... Zur Vermeidung von Informationsüberflutung der Testpersonen, sie gegenüber der Untersuchung nicht zu demotivieren oder Hürden für die Informationsgenerierung aufzubauen, wurde eine tägliche Maximalgrenze für das Vornehmen von Tagebucheinträgen festgelegt. Die Probanden mussten somit nicht mehr als zwei bis drei Arbeitsaufgaben pro Tag im Tagebuch notieren.

Um den Aufwand der Tagebuchführung für die Testpersonen gering zu halten und so möglichst viele relevante Daten generieren zu können, stand das Tagebuch in verschiedenen technologischen Formen zur Verfügung:

- Vordrucktes Formular, auch als *Paper-and-Pencil* bezeichnet,
- elektronisches Formular zum Versenden per E-Mail,
- elektronisches Online-Formular,
- verbal per Telefonanruf an den Studienleiter.

Ein weiterer Vorteil, der multitechnologischen Umsetzung des Tagebuchs, war eine präzise Verfolgung der Tagebuchführung in Echtzeit. So konnte die Nutzung der Testsoftware nachvollzogen und gesteuert werden. Bei längeren Lücken in der Tagebuchführung eines Probanden erfolgte eine kurze Nachfrage zur Nutzung der Testsoftware über den beobachteten Zeitraum und falls notwendig, eine Erinnerung zur Führung des Tagebuchs.

Zusätzlich zu den Erhebungsinstrumenten beinhaltete das Tagebuch (vgl. Anhang C) noch Abfragen zum Namen der Testperson, Datum und Uhrzeit der Eintragung, Dauer der Aufgabe, Verwendung von Interaktionselementen und Beschreibung der Aufgabe. Der Name der Testperson diente der Zuordnung des Tagebucheintrags zu einem Probanden. Die weiteren Daten gaben Aufschluss zu Rahmenbedingungen der Nutzung, konnten ggf. für die Erklärungsfindung individueller Nutzungssituationen und zur Analyse des Nutzerverhaltens herangezogen werden. Angaben zum Namen der Testperson, Datum und Uhrzeit sowie der Beschreibung der Aufgabe wurden durch offene Fragen erhoben. Die Abfrage nach der Verwendung von Interaktionselementen erfolgte entsprechend der Beschreibung in Kapitel 3.3.1. Für die Dauer der Aufgabe gab es folgende vorgegebene Wahlmöglichkeiten:

- Etwa 1 Stunde und länger,
- über 15 Minuten, aber unter 1 Stunde,
- unter 15 Minuten.

Das Tagebuch enthielt außerdem einen Abschnitt mit Erläuterungen zur Tagebuchführung entsprechend der Beschreibungen in diesem Kapitel und der im Tagebuch verwendeten Erhebungsinstrumente.

### **3.3.3 Interventionen**

#### **3.3.3.1 Vor-Ort-Beobachtung**

Während der Vor-Ort-Beobachtungen waren die Testpersonen angehalten, in Anwesenheit des Studienleiters, in ihrem gewohnten Arbeitsumfeld eine ihrer normalen Aufgaben mit der Testsoftware zu bearbeiten. Dabei sollten die Probanden möglichst wenig durch die Anwesenheit des Studienleiters beeinflusst werden. So gab es keine Erläuterungen der Probanden zu der Aufgabenbearbeitung und auch keine unmittelbaren Nachfragen zu Nutzungssituationen. Ziel der Vor-Ort-Beobachtungen war es, als Ergänzung zu den Daten der subjektiven Einschätzungen der Probanden, stichprobenartig objektive Daten zum Nutzerverhalten zu gewinnen. Die Datenerhebung erfolgte sowohl quantitativ als auch qualitativ.

Die quantitative Datenerhebung geschah durch die A/V-Aufzeichnung der Bildschirmaktivitäten mittels einer *ScreenCam*, einer Software, die als Hintergrundprozess alle Aktivitäten auf dem Bildschirm aufzeichnet und in einer Videodatei abspeichert (Fast, 2002). Vorteile einer ScreenCam gegenüber Aufzeichnungen durch Videorecordern und -kameras sind die unkomplizierte technische Handhabung, die Vollständigkeit der A/V-Aufnahme des Bildschirms und zusätzliche Features, wie die Aufzeichnung von Mausklicks. Für die Vor-Ort-Beobachtung wurde die ScreenCam-Software auf dem PC des Probanden installiert und entsprechend der folgenden Einstellungen konfiguriert:

- Aufzeichnung des kompletten Bildschirms,
- Samplerate bei 10 Bildern pro Sekunde,
- Aufzeichnung des Mauspfeils,
- Aufzeichnung linker und rechter Mausklicks durch unterschiedlich farbige Kreise,
- kein akustischer Mitschnitt.

Eine Aufzeichnungsdatei umfasste immer eine Aufgabenbearbeitung.

Zeitgleich mit der A/V-Aufzeichnung zur Aufgabenbearbeitung der Probanden, erfolgte eine Protokollierung durch den Studienleiter unter Zuhilfenahme eines teilstrukturierten Beobachtungsprotokolls (vgl. Anhang D). Das Protokoll unterteilte sich in drei Abschnitte.

- Der erste Abschnitt beinhaltete den Namen der Testperson, Datum und Uhrzeit der Erstellung, Dauer der Aufgabe, Verwendung von Interaktionselementen. Die Notierung sämtlicher Daten erfolgte entsprechend der Ausführung in Kapitel 3.3.2.
- Allgemeine Beobachtungen zum Nutzerverhalten wurden im zweiten unstrukturierten Abschnitt anhand von Notizen festgehalten.
- Der letzte Abschnitt des Protokolls diente der Aufzeichnung von Beobachtungen zu Interaktionen, speziell zu „Usability-Problemen“ und „Unterstützenden Effekten“. Dementsprechend enthalten waren die Uhrzeit des beobachteten „Usability-Problems“ oder „Unterstützenden Effekts“, eine Liste mit den definierten „Usability-Problemen“ und „Unterstützenden Effekten“ (vgl. Kapitel 3.2.3) zur Markierung des beobachteten Ereignisses und Beobachtungen zu den Interaktionen.

Im Anschluss an die eigentliche Vor-Ort-Beobachtung machten die Probanden wie gewohnt einen Eintrag in das Tagebuch für die bearbeitete Aufgabe.

### **3.3.3.2 Fragebogen zur Intervention**

Im zweiten Teil der Interventionen bekamen die Teilnehmer einen Fragebogen als Vordruck (vgl. Anhang E) ausgehändigt. Anhand der gestellten Fragen sollten sie ihre Eindrücke und Einschätzungen zur Aufgabenbearbeitung über den vergangenen Studienzeitraum abgeben.

Der Fragebogen bestand aus einem Abschnitt mit einer kurzen Erläuterung zur Handhabung, Name der Person und Datum für die Zuordnung des Fragebogens sowie einem Abschnitt mit den Erhebungsinstrumenten. Am Ende erfolgte eine Danksagung.

### **3.3.3.3 Recall Survey**

Recall Surveys kamen in der Studie nur optional zum Einsatz und sollten klärungsbedürftige Punkte der Datenerhebung auflösen. Vor jeder Intervention wurden die Tagebucheinträge des vergangenen Nutzungsabschnitts einer Testperson untersucht, die Ergebnisse und zu klärende Punkte in einem Recall Survey Protokoll vermerkt (vgl. Anhang F). Dazu war das Recall

Survey Protokoll in zwei Abschnitte gegliedert. Der erste Abschnitt diente der Zuordnung des Protokolls und beinhaltete Angaben zu

- Zählung der Intervention,
- Name der Person,
- Zuordnung in die Adaptive oder Nicht-Adaptive Gruppe,
- Daten von Beginn und Ende des Nutzungsabschnitts.

Im zweiten Abschnitt wurden die Ergebnisse der Tagebuchsichtung eingetragen. Ergaben sich daraus klärungsbedürftige Punkte bekamen die betreffenden Ergebnisse einen Vermerk. Die Probanden wurden dann während der Intervention zu den Anmerkungen befragt und das Protokoll durch die Aussagen notizenartig ergänzt.

### **3.3.4 Dokumentation der Benutzungsoberfläche**

Die Dokumentation der Benutzungsoberfläche erfolgte mittels sogenannter *Screenshots* (vgl. Anhang G), als Bilddateien gespeicherten visuellen Momentaufnahmen eines Bildschirms (Zimmermann, 2004).

Ein vor Studienbeginn angefertigter Screenshot der Benutzungsoberfläche der komplett neu installierten Software diente als Grundlage für die Dokumentation.

Bereits mit dem eigentlichen Beginn der Studie, in der Einweisung der Studienteilnehmer (vgl. Kapitel 3.4.3), wurden die ersten Screenshots angefertigt. Sie dokumentierten den Ausgangszustand des UI bei dem jeweiligen Probanden.

Dann erfolgten die Screenshots während jeder Intervention und bei der Abschlussbefragung.

### **3.3.5 Abschlussbefragung**

#### **3.3.5.1 Fragebogen zum Studienabschluss**

Der Fragebogen zum Studienabschluss entsprach in Aufbau und Vorlage dem Fragebogen zur Intervention (vgl. Kapitel 3.3.3.2) plus einem Abschnitt zu Verbesserungspotentialen (vgl. Anhang H). Der zusätzliche Abschnitt bestand aus den Erhebungsinstrumenten zur Einstellung gegenüber Formen der Personalisierung (vgl. Kapitel 3.2.1) und den speziell für die Abschlussbefragung ergänzten Anforderungen an adaptive Systeme (vgl. Kapitel 3.2.4).

Sämtliche Formulierungen wurden entsprechend der Fragestellung zur abschließenden Bewertung angepasst.

### **3.3.5.2 Interview zum Studienabschluss**

Im zweiten Teil der Abschlussbefragung erfolgte ein teilstrukturiertes Interview. Es fokussierte vor allem die Arbeitsweise der Probanden sowie die Wahrnehmung der Adaptivität und die Motivation zur Personalisierung. Die Durchführung geschah auf der Basis eines Interviewprotokolls (vgl. Anhang I), welches sich wie folgt gliederte:

- Zuordnung des Protokolls und unabhängige Variablen der Testperson,
- Fragen zur Arbeitsweise,
- Fragen zu Personalisierung und Adaptivität,
- Zusammenfassung von Bewertungen der Testperson im Studienverlauf.

Die Angaben im ersten Abschnitt wurden bereits vor dem Interview als frei formulierter Text eingetragen, es handelte sich um:

- Den Namen der Testperson,
- die Zugehörigkeit zur Adaptiven oder Nicht-Adaptiven Gruppe,
- das Datum der Abschlussbefragung,
- die Angaben der Probanden aus dem Fragebogen zum Studienbeginn zu Nutzungserfahrungen, Einstellung gegenüber Formen der Personalisierung, Nutzungshäufigkeit und Tätigkeitsbeschreibung,
- die während des Studienverlaufs gewonnen Einschätzungen des Testleiters zu Nutzungserfahrungen und Nutzungshäufigkeit der Probanden.

Die Fragen zur Arbeitsweise im zweiten Abschnitt gliederten sich in zwei Teile:

- Der erste Teil beinhaltete vom Studienleiter gewonnene Eindrücke zur Arbeitsweise des Probanden in einem frei formulierten Text.
- Im zweiten Teil sollte der Teilnehmer seine Arbeitsweise und Eindrücke zu verwendeten Interaktionselementen der Testsoftware schildern. Er wurde aufgefordert zu erläutern auf welche Interaktionselemente er in welcher Nutzungssituation zurückgreift, was ihn an den jeweiligen Interaktionselementen gefällt oder stört und welche Beobachtungen er während der Nutzung der Interaktionselements gemacht hat.

Abschnitt drei setzte sich eingehender mit dem Thema Personalisierung und Adaptivität auseinander. Ein erster Vermerk gab an mit welcher Einstellung die Testperson vor der Studie gearbeitet hatte, adaptiv oder nicht-adaptiv oder ob das Testprodukt erst für die Studie installiert wurde. Die zweite Fragestellung beschäftigte sich mit der Wahrnehmung der adaptiven Anpassungen während der Studie, sie richtete sich somit auch nur an die Probanden der Adaptiven Gruppe. Diese sollten dann beschreiben, ob sie überhaupt Anpassungen während der Studie wahrgenommen hatten, wenn ja, ob sie diese in Verbindung mit den Fragestellungen zu den „Ereignissen“ in den Erhebungsinstrumenten brachten und ob sie diese als Reflektion ihrer Arbeitsweise wahrnahmen. Die folgenden Fragen richteten sich wieder an beide Testgruppen und adressierte den Wunsch und die Gründe bestimmte Interaktionselemente der Testsoftware personalisieren oder nicht personalisieren zu wollen. Anschließend sollten sie erläutern welche konkrete Form der Personalisierung sie bei den angegeben Interaktionselementen der Testsoftware kennen und welche sie vornehmen würden<sup>5</sup>. Bei allen Fragestellungen wurden bereits durchgeführte Personalisierungen vor oder während der Studie berücksichtigt.

Im vierten Abschnitt wurden die Ergebnisse einer ersten Vorab-Auswertung der erhobenen Daten zusammengefasst. Mögliche Fragen für das Interview befanden sich in den Feldern „Fragen zur Bewertung an den Studienteilnehmer“ und „Weitere Vorkommnisse“, in welche die Antworten der Probanden dann ebenfalls notiert wurden.

### **3.4     *Durchführung der Studie***

#### **3.4.1    Auswahl des Testprodukts**

Als Testprodukt für die Durchführung der Studie diene Microsoft Word XP. Die Auswahl wurde aufgrund der Tatsache getroffen, dass bei Microsoft Word Adaptivität bei Interaktionselementen bereits in der zweiten Generation umgesetzt wird und eine Vielzahl von Anwendern auf dieses Software-Produkt für die Erledigung ihrer Aufgaben in Beruf und Studium zurückgreifen.

---

<sup>5</sup> Bei der gewählten Testsoftware Microsoft Word XP (vgl. Kapitel 3.4.1) sind nicht alle Interaktionselemente gleich personalisierbar. Für die Menüs besteht nur die Möglichkeit adaptive Anpassungen zu aktivieren oder zu deaktivieren, entsprechend den Einstellung der Studie für die beiden Testgruppen. Die Symbole können adaptiv verwendet werden, sind gleichzeitig auch adaptierbar. Die Adaptierbarkeit der Symbole bezieht sich sowohl auf einzelne Symbole als auch auf Symbolblöcke zu bestimmten Funktionskategorien, welche in der Regel aus ca. 10 bis 20 Symbolen bestehen. Die Tastatur-Shortcuts sind individuell adaptierbar.



Die von Microsoft verwendete Technologie zur Adaptivität nennt sich MS IntelliSense und kommt in der Anpassung der Menüs und Symbole sowie dem Assistenten zum Einsatz. Aufgrund des Pilotstudiencharakters der Untersuchung, beschränkte sich die Arbeit auf die adaptiven Menüs, den sogenannten Smart Menus, die ebenfalls adaptiven Symbole wurden nicht berücksichtigt. Tests vor Beginn der Studie lieferten folgende für die Untersuchung relevanten Ergebnisse bezüglich der adaptiven Funktionalität von MS IntelliSense bei den Menüs von Microsoft Word XP:

- Die adaptiven Menüs von Microsoft Word XP gliedern sich in einen sichtbaren Teil mit den genutzten Menüeinträgen und einen verdeckten Teil mit den nicht genutzten Menüeinträgen.
- Um den verdeckten Teil sichtbar zu machen, hat der Nutzer die Möglichkeit diesen manuell aufzuklappen oder er stellt eine Option ein, die den verdeckten Teil nach ca. 3 bis 4 Sekunden automatisch öffnet.
- Es besteht die Möglichkeit die Gesamtzahl adaptiver Anpassungen zurückzusetzen ohne die Software neu installieren zu müssen, einzelne Anpassungen können manuell nicht zurückgesetzt werden.
- Adaptive Anpassungen bei den Menüs von Microsoft Word XP gibt es als Verschwinden von Menüeinträgen aus dem sichtbaren in den verdeckten Teil der Menüs, dem Verschieben von Menüeinträgen aus dem verdeckten in den sichtbaren Teil und als situationsabhängige Anpassungen z.B. bei der Bearbeitung von Grafiken und Objekten.
- Das Verschwinden eines Menüeintrags aus dem sichtbaren Teil richtet sich nach der Nutzungshäufigkeit der betreffenden Menüliste.
- Ein Menüeintrag verschiebt sich bereits nach einmaliger Verwendung vom verdeckten in den sichtbaren Teil.
- Einzelne Einträge der Menüs ordnen sich bei der adaptiven Anpassung nicht entsprechend der Nutzungshäufigkeit, sondern nach der standardmäßigen Reihenfolge der Einträge in der Menüliste.

### **3.4.2 Anwerbung und Auswahl der Studienteilnehmer**

Die Anwerbung der Studienteilnehmer erfolgte mittels einer Bewerberliste für Usability-Tests der Abteilung CT IC7 und durch Neurekrutierung an Hochschulen, Studentenwohnheimen,

Bibliotheken und Siemens-Standorten. Für die Teilnahme wurde vorausgesetzt, dass mögliche Testpersonen über einen eigenen Computer verfügten und im Raum München wohnten oder arbeiteten.

Die in der Liste aufgeführten Personen wurden per E-Mail kontaktiert und zur Studie eingeladen. An die E-Mail angehängt befand sich eine Datei, die den Kontaktierten einen ersten Eindruck zu den Voraussetzungen einer Studienteilnahme, zum Studienablauf, dem geplanten Termin und ihrer Vergütung verschaffen sollte (vgl. Anhang A). Bei Interesse wurden die Personen um Rückmeldung gebeten, Kontaktdaten waren sowohl in der E-Mail als auch im Dateianhang angegeben. Die Neurekrutierung erfolgte durch das Aufhängen von Plakaten und dem Auslegen von Flugzetteln. Beide beinhalteten die gleichen Informationen wie der beschriebene Dateianhang und waren grafisch identisch gestaltet. Für die Formulierungen der Dateianhänge, Plakate und Flugzettel wurde zwischen Siemens-Mitarbeitern und Siemens-externen Personen unterschieden. Dies geschah aufgrund der unterschiedlichen Vergütung und vor allem um gegenüber beiden Gruppen seriös aufzutreten. Auf die eigentliche Auswahl der Teilnehmer und den weiteren Verlauf der Studie hatte diese Einteilung keinen Einfluss.

Meldeten sich Personen auf die E-Mails oder Aushänge, bekamen sie den Fragebogen zum Studienbeginn (vgl. Anhang B), mit der Bitte, diesen ausgefüllt an die angegebene Kontaktadresse zurückzuschicken.

Ausschlusskriterium für eine Studienteilnahme war die Angabe, nicht mit den Menüs von Microsoft Word XP zu arbeiten, da es sich hier um das zentrale Untersuchungsobjekt handelte. Für die Teilnehmerauswahl diente die Erhebung der

- Zuordnung in eine Nutzergruppe aufgrund der Nutzungserfahrungen,
- Nutzungshäufigkeit des Testprodukts,
- Einstellungen gegenüber Formen der Personalisierung.

Nicht als Auswahlkriterium herangezogen wurden die Demographie und die Tätigkeitsbeschreibung der Nutzer, weil beide für die Beantwortung der Fragestellungen letztlich nicht relevant waren.

Da bei der Zuordnung in eine Nutzergruppe ein großer Teil der Probanden nicht eindeutig, entsprechend der Einteilung von Burmester in Novizen- bzw. Expertennutzer (Burmester, 2003) kategorisiert werden konnte, erfolgt die Zuordnung in die drei folgenden Nutzergruppen:

- Novizen mit einem sehr geringen Kenntnisstand,
- User mit mittlerem Kenntnisstand entsprechend einem durchschnittlichen Anwender,

- Experten mit sehr umfangreichen Kenntnissen, speziell auch bezüglich der Funktionsweise und Usability von EDV-Systemen.

Die Gliederung der Bewerber nach ihrer Einstellung zu Formen der Personalisierung teilte sich in vier Kategorien möglicher personalisierter Systeme entsprechend der in Kapitel 3.2.1 erläuterten Ausprägungen auf. Die gebildeten Gruppen entsprachen dem Wunsch der Personen nach einem

- „Adaptiven UI“ aufgrund des Wunsches nach automatischen Anpassungen.
- Gab die Person an, sie wünscht sich automatische Anpassungen mit völliger oder weitgehender Systemautonomie plus der Möglichkeit auch selbst Anpassungen vornehmen zu können, dann war dies gleichbedeutend mit der Definition eines „Mixed-Initiative UI“.
- Der Wunsch selbst Anpassungen vornehmen zu können, entsprach der Kategorie eines „Adaptierbaren UI“. Eine Kombination mit den Wunsch nach automatischen Anpassungen mit kontrollierter Systemautonomie fiel ebenfalls in diese Kategorie, da davon ausgegangen wurde, dass die Person die automatischen Anpassungen nur als Entscheidungshilfe zur Adaptierung verwendet.
- Der Wunsch nach einer Standardversion entsprach der Kategorie des „Nicht personalisierbaren UI“.

Für die Studie waren 16 Teilnehmer vorgesehen. Diese Anzahl schien sowohl organisatorisch als auch für die Generierung von ersten relevanten Ergebnissen durch eine Pilotstudie realistisch. Die Erfüllung von interner und externer Validität, eine repräsentative Abbildung der Population aller möglichen Nutzer von Microsoft Word XP oder von interaktions-adaptiven Systemen im allgemeinen war dadurch allerdings nicht gegeben. Eine sinnvolle Alternative bot die Bildung einer möglichst heterogenen Stichprobe auf der Basis der festgelegten Auswahlkriterien, um so unterschiedliche Nutzer mit der Adaptivität konfrontieren zu können. Auf die Studie bewarben sich insgesamt 28 Personen mit folgenden, in der Matrix von Tabelle 1 dargestellten, Merkmalsausprägungen der berücksichtigten drei Kriterien für die Teilnehmerauswahl:

Einstellung gegenüber Formen der Personalisierung	Adaptives UI	Mixed-Initiative UI	Adaptierbares UI	Nicht personalisierbares UI
Nutzungshäufigkeit				
Täglich, über 4 Stunden			1 Experte 1 User	
Täglich, aber unter 4 Stunden	1 Experte 3 User 1 Novize	2 User 1 Novize	2 Experten 1 User 1 Novize	2 Experten
Mehrmals wöchentlich	1 User 1 Novize	1 Experte 2 User	1 Experte 2 User	2 User
Nicht mehr als einmal pro Woche	1 Novize		1 User	

**Tabelle 1 - Matrix der Studienbewerber**

Für die Studie ergab sich letztendlich die folgende Stichprobe:

Einstellung gegenüber Formen der Personalisierung	Adaptives UI	Mixed-Initiative UI	Adaptierbares UI	Nicht personalisierbares UI
Nutzungshäufigkeit				
Täglich, über 4 Stunden			1 Experte 1 User	
Täglich, aber unter 4 Stunden	1 User 1 Novize	1 User 1 Novize	1 Novize	1 Experte
Mehrmals wöchentlich	1 User	1 Experte 1 User	2 User	2 User
Nicht mehr als einmal pro Woche	1 Novize			

**Tabelle 2 - Matrix der Studienteilnehmer**

Die Zuordnung der Probanden zur Adaptiven und Nicht-Adaptiven Gruppe erfolgte gleichmäßig entsprechend der Kriterien:

Einstellung gegenüber Formen der Personalisierung	Adaptives UI	Mixed-Initiative UI	Adaptierbares UI	Nicht personalisierbares UI
Nutzungshäufigkeit				
Täglich, über 4 Stunden			1 Experte Adaptiv  1 User: Nicht Adaptiv	
Täglich, aber unter 4 Stunden	1 User: Adaptiv  1 Novize: Nicht Adaptiv	1 User: Nicht Adaptiv  1 Novize: Adaptiv	1 Novize: Nicht Adaptiv	1 Experte: Adaptiv
Mehrmals wöchentlich	1 User: Nicht Adaptiv	1 Experte: Nicht Adaptiv  1 User Adaptiv	1 User: Adaptiv  1 User: Nicht Adaptiv	1 User: Adaptiv  1 User: Nicht Adaptiv
Nicht mehr als einmal pro Woche	1 Novize: Adaptiv			

**Tabelle 3 - Matrix der Testgruppen**

Nach der Auswahl wurden die ausgesuchten Studienteilnehmer erneut kontaktiert und ein Termin für die Einweisung ausgemacht.

### 3.4.3 Einweisung der Studienteilnehmer

Die Einweisung der Studienteilnehmer fand vor Ort bei den Probanden statt. Ziel war es, den Teilnehmern den Studienablauf inhaltlich und terminlich detailliert zu erläutern. Außerdem markierte sie gleichzeitig den individuellen Studienbeginn für den jeweiligen Probanden.

Die Erläuterungen zur Studie erfolgten auf der Basis eines Leitfadens zur Studie, welcher den Teilnehmern ausgehändigt wurde. Einleitend wurde kurz geschildert, dass sich die Abteilung CT IC7 mit der benutzerfreundlichen Gestaltung von Bedienkonzepten beschäftigt und die Studie auch dieses Ziel verfolgt. Eine Aufklärung über die genauen Studienfragen zum Thema Personalisierung und Adaptivität gab es jedoch nicht, die Probanden sollten sich nicht von Anfang an zu stark auf diesen Aspekt fokussieren. Geplant war vielmehr eine bewusste Heranführung an die Thematik. Die Untersuchung erhielt deshalb formal die Bezeichnung „Studie zur Ergonomie von Office-Software“. Die Teilnehmer wurden auch nur darüber

informiert, dass für den Studienzeitraum spezifische Einstellungen an ihrer Software vorgenommen werden, die sie nicht verändern sollten. Der zweite Abschnitt schilderte den Studienablauf. Neben der Studiendauer und dem geplanten Zeitraum waren dort auch detaillierte Angaben zum Ablauf der Interventionen, zum Zweck des Tagebuchs und zur Abschlussbefragung untergebracht. Abschnitt drei beschrieb die Handhabung des Tagebuchs entsprechend den Ausführungen in Kapitel 3.3.2. Um einen Überblick über die Tagebuchführung der Teilnehmer zu erhalten, wurde nach Präferenzen der technologischen Formen gefragt und diese notiert. Außerdem erfolgte eine exemplarische Demonstration der Tagebuchführung anhand eines Testeintrags in das Online-Tagebuch. Im nächsten Abschnitt wurde anhand eines Kalenderausschnitts der Ablauf der Studie skizziert. Eine Verpflichtung zum Datenschutz von Seiten der Siemens AG, eine Teilnahmeerklärung der Probanden und Kontaktdaten für Rückfragen befanden sich im letzten Abschnitt des Leitfadens.

Vor dem Start der Studie wurde festgehalten, ob die Testperson bislang mit adaptiven oder nicht-adaptiven Menüs arbeitete oder ob das Testprodukt Microsoft Word XP erst neu installiert war. Dann wurden die adaptiven Anpassungen zurückgesetzt und entsprechend der Testgruppenzugehörigkeit die Adaptivität aktiviert bzw. deaktiviert, für die Adaptivität wurde zusätzlich das Setup zur manuellen Verlängerung der Menüs eingestellt.

Es erfolgten die ersten Screenshots zur Dokumentation der Benutzungsoberfläche.

Am Ende der Einweisung stand die Klärung noch offener Fragen der Testpersonen, die Vereinbarung eines Termins für die erste Intervention und die Bitte Urlaube und längere Abwesenheiten zwecks der Studienkoordinierung mitzuteilen.

#### **3.4.4 Ablauf der Studie**

Das Studienkonzept sah eine Studiendauer von neun Wochen mit drei Interventionen pro Teilnehmer vor, die Abschlussbefragung beschloss die Studie. Da die Studie als Feldstudie konzipiert war und alle Treffen vor Ort bei den Teilnehmern stattfanden, kam es bei einigen Probanden zu terminlichen Engpässen aufgrund von Urlaubswochen oder geschäftlichen Verpflichtungen, so dass die eine Mindeststudiendauer von acht Wochen mit zwei Interventionen festgelegt wurde.

### 3.5 *Quantitative Auswertung*

Bei den, durch die Erhebungsinstrumente (vgl. Kapitel 3.2.1 - 3.2.7), gewonnenen Daten der subjektiven Einschätzung handelt es sich um zwei Arten von Daten:

- Intervallskalierte Messdaten aus Ratingskalen<sup>6</sup>,
- Häufigkeitsziffern aus Auszählungen bei der Anzahl von „Ereignissen“ und der Einstellung gegenüber Formen der Adaptivität.

Die Auswertung erfolgte in einem ersten Schritt deskriptiv durch die Berechnung von Mittelwerten und Standardabweichung für alle Daten.

Um zu beurteilen, ob die Verhältnisse der deskriptiven Ergebnisse auf Zufall rückführbar oder generalisierbar waren, wurden in einem zweiten Schritt Signifikanztests für die Erhebungen mit einem großen Stichprobenumfang aus dem Tagebuch (vgl. Kapitel 3.3.2) durchgeführt. Da für die Stichproben die Populationscharakteristika nicht bekannt waren, wurden die Daten verteilungsfrei auf ihre Signifikanz hin getestet. Für den Vergleich der Daten über die drei Nutzergruppen in der fünften Fragestellung erfolgte der Mediantest für k unabhängige Stichproben (Bortz, Lienert, 2003, S.134-137), sonst wurde der U-Test für zwei unabhängige Stichproben nach Mann-Whitney (Bortz, Lienert, 2003, S.138-147) verwendet. Als Signifikanzniveau wurde der gängigen Standardwert von  $\alpha < 0,05$  (Bortz, Lienert, 2003, S.37) gewählt.

Eine Ausnahme bildeten die gewonnenen Daten des AttrakDiff, da es sich hier um kein speziell für die Studie entwickeltes Erhebungsinstrument handelte. Die Auswertung erfolgte entsprechend der standardmäßig gültigen Vorgaben (AttrakDiff, 2004).

Zur Beurteilung, ob die Probanden mentale Anstrengung von inhaltlichen Einflüssen der Aufgabenbearbeitung unterscheiden konnten, wurde für die in Kapitel 3.2.2 beschriebenen Erhebungsinstrumente BSMA und Aufgabenschwierigkeit der Korrelationskoeffizient (Bortz, 1989, S.248-287) nach Spearman ermittelt. Es sollte festgestellt werden, ob die Ergebnisse des BSMA mit den Angaben der Aufgabenschwierigkeit korrelieren. Eine Korrelation von über 0,5 würde vermuten lassen, dass die Probanden beide Fragestellungen letztlich nicht völlig voneinander trennen konnten.

---

<sup>6</sup> Laut Bortz, Döring handelt es sich bei Messdaten aus Ratingskalen um Daten, die als intervallskaliert interpretiert werden können (Bortz, Döring, S.175).

Die Auswertung der objektiven quantitativen Daten (vgl. Kapitel 3.2.9) erfolgte durch Auszählung.

Bei der Dokumentation der Benutzungsoberfläche wurde das Aussehen des Interaktionselements über den Studienzeitraum von Nutzungsabschnitt zu Nutzungsabschnitt verglichen. Festgehalten wurden die adaptiven Anpassungen in ihrer Gesamtzahl, bestehend aus Menüeinträgen, die aus dem sichtbaren in den verdeckten Teil der Menüs verschwanden und Menüeinträgen, die aus dem verdeckten in den sichtbaren Teil verschoben wurden.

Die bei der A/V-Aufzeichnung gewonnenen Daten bestanden aus Werten zu den Arbeitsschritten und Bearbeitungszeiten einer Interaktion. Für jede Interaktion wurde ein Vergleich zwischen tatsächlicher und optimaler Anzahl von Bearbeitungsschritten angestellt. Es ergab sich eine Anzahl nicht optimaler Interaktionen, bei denen zusätzliche Arbeitsschritte zur Komplettierung benötigt wurden.

### **3.6 Qualitative Auswertung**

Eine Auswertung, der in Kapitel 3.2.8 vorgestellten qualitativen Daten, fand anhand der in Bortz, Döring beschriebenen Arbeitsschritte statt (Bortz, Döring, 2003, S.329-331). Für die Beantwortung der Studienfragen hinzugezogen wurden die folgenden Kategorien:

- Allgemeine Gebrauchstauglichkeit der Testsoftware,
- Einstellung gegenüber Formen der Personalisierung,
- subjektive Einschätzung der Adaptivität,
- Beobachtungen zu Nutzungserfahrungen der Probanden,
- Beobachtungen zur Wahrnehmung der Probanden.

Zur Evaluation der Erhebungsinstrumente wurde Feedback der Studienteilnehmer zum Studienablauf und den Erhebungsinstrumenten erhoben. Zu folgenden Themen konnten Daten gesammelt werden:

- Allgemeine Daten zum Studienablauf,
- Bewertung im Mittel in den Fragebogen zur Intervention, durchgeführt bei den Erhebungsinstrumenten BSMA und Anforderungen an adaptive Systeme,
- Erhebungsinstrument Anforderungen an adaptive Systeme,
- Erhebungsinstrument AttrakDiff,
- Erhebungsinstrument „Usability-Probleme“ und „Unterstützende Effekte“,
- Führung des Tagebuchs.



Alle Daten wurden anhand ihrer Kennzeichen nach einer positiven oder negativen subjektiven Einschätzung geordnet. Zur Bewertung herangezogen wurden Gesamttendenzen und besonders relevant erscheinende Einzelfälle.

### **3.7    *Tücken der Datenerhebung***

Für die Datenerhebung des BSMA erwies es sich als tückisch, das Tagebuch multitechnologisch bereitgestellt zu haben. Die als Paper-and-Pencil konstruierte 220-stufige Skala konnte in den elektronischen Versionen des Tagebuchs nicht abgebildet werden. Für die Online-Version ergaben sich 19 darstellbare Abstufungen der Bewertung, für das elektronische E-Mail-Formular 23 (vgl. Anhang J). Die gewonnenen Werte mussten dann auf die 220-stufige Skala umgerechnet werden.

Bei dem Erhebungsinstrument Anforderungen an adaptive Systeme mussten ebenfalls Daten an die 5-stufige Ratingskala umgerechnet werden.

Zum einen wurde, wie in Kapitel 3.2.4 beschrieben, die Dimension „Nutzen“ nicht im Tagebuch erhoben, sondern aufgrund der Werte zur Intensität von „Ereignissen“ (vgl. Kapitel 3.2.3) berechnet. Da diese Daten ebenfalls aus 5-stufigen Ratingskalen resultierten, passten sich die Ergebnisse nicht exakt in die Skalierung der Werte der Anforderungen an adaptive Systeme an. Um routinemäßigen Bewertungen der Dimensionen der Anforderungen an adaptive Systeme gegenzusteuern, wurden in den ausgehändigten Vordrucken von Tagebuch und Fragenbogen zur Intervention und zum Studienabschluss auch teilweise 7-stufige Ratingskalen eingesetzt. Auch diese gewonnenen Werte mussten in die 5-stufige Skala umgerechnet werden.

Zu Beginn der Studie hatten drei der Probanden Verständnisprobleme mit den im Tagebuch gestellten Fragen, einige Tagebucheinträge konnten deshalb nicht ausgewertet werden. Bei zwei dieser Teilnehmer führte dies zu je einem Nutzungsabschnitt ohne Tagebucheintrag.

Außerdem gab es bei einem Probanden einen Nutzungsabschnitt, in dem sich kein vorgenommener Tagebucheintrag auf die Nutzung des untersuchungsrelevanten Interaktionselements bezog.

Ein weiterer Studienteilnehmer führte in zwei Nutzungsabschnitten keine Arbeitsaufgabe mit der Testsoftware Microsoft Word XP aus. Diese Nutzungsabschnitte beinhalteten dadurch keine auswertbaren Tagebucheinträge.

Bei vier Testpersonen konnte nur die Mindestanzahl von zwei Interventionen durchgeführt werden.

Zudem erwiesen sich die von der ScreenCam-Software gestellten Mindestanforderungen an ein System als diffizil. Die PCs einiger Studienteilnehmer konnten diesen Anforderungen nicht gerecht werden, ein einwandfreies Arbeiten während der Vor-Ort-Beobachtungen mit laufender Aufzeichnung war nicht möglich. Bei drei Probanden wurde aus diesem Grund auf die A/V-Aufzeichnung verzichtet, von den Aufzeichnungen der übrigen Teilnehmer wurden drei weitere beim Datentransfer beschädigt und waren nicht auswertbar.

Ein Proband stellte im Verlauf der Studie das System von adaptiv auf nicht-adaptiv um. Er bezog diese Umstellung jedoch nicht auf die für die Studie gewählten Einstellungen, so dass der genaue Termin der Änderung nicht erfasst werden konnte. Der betreffende Nutzungsabschnitt des Probanden floss deshalb nicht in die Bewertung ein.

Die Auswertung der Kategorien Nutzungshäufigkeit und Nutzungserfahrungen der Probanden ergab eine veränderte Matrix der Auswahlkriterien:

Einstellung gegenüber Formen der Personalisierung	Adaptives UI	Mixed-Initiative UI	Adaptierbares UI	Nicht personalisierbares UI
Nutzungshäufigkeit				
Täglich, über 4 Stunden	User: Adaptiv		Experte: Nicht Adaptiv	
Täglich, aber unter 4 Stunden			Novize: Nicht Adaptiv	
Mehrmals wöchentlich		Experte: Nicht Adaptiv  User: Adaptiv  Novize: Adaptiv	Experte: Adaptiv  User: Adaptiv  User: Nicht Adaptiv	
Nicht mehr als einmal pro Woche	User: Nicht Adaptiv  Novize: Adaptiv  Novize: Nicht Adaptiv	User: Nicht Adaptiv		Experte: Adaptiv  User: Adaptiv  Novize: Nicht Adaptiv

**Tabelle 4 - Matrix der Studienteilnehmer aufgrund der Auswertung der qualitativen Daten**

Durch diese Veränderungen lieferten die Untersuchungen nicht so weit gefächerte Ergebnisse zu Nutzern mit unterschiedlicher Nutzungshäufigkeit wie ursprünglich geplant. Es fehlten Anwender, die täglich mit der Testsoftware arbeiteten.

## 4 Ergebnisse

### Erste Forschungsfrage

Die erste Forschungsfrage beschäftigte sich mit dem Auftreten und den Auswirkungen von Usability-Problemen und unterstützenden Effekten. Es sollten Unterschiede bei den Parametern Anzahl, Häufigkeit und Intensität der Erhebungsinstrumente „Usability-Probleme“ und „Unterstützenden Effekte“ (vgl. Kapitel 3.2.3) zwischen Adaptiver und Nicht-Adaptiver Gruppe identifiziert und lokalisiert werden. Außerdem untersuchte die Forschungsfrage, ob sich ein längerer Nutzungszeitraum bei der Nutzung adaptiver Software positiv auf die Parameter Anzahl, geschätzte Häufigkeit und Intensität von „Usability-Problemen“ auswirkt.

Für eine genauere individuelle Analyse der gewonnenen subjektiven Daten wurden die objektiven quantitativen Daten aus den Vor-Ort-Beobachtungen herangezogen

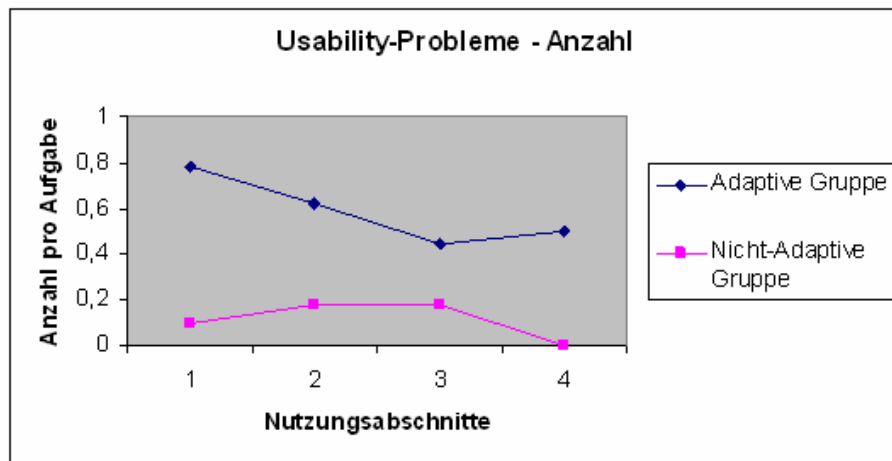
Bei der Betrachtung des Studienzeitraums ergab sich, aufgrund der gemachten Tagebucheinträge der Probanden in den Nutzungsabschnitten, der in Tabelle 5 dargestellte Umfang an Stichproben für die Testgruppen:

Studienzeitraum	Nutzungs- abschnitt 1	Nutzungs- abschnitt 2	Nutzungs- abschnitt 3	Nutzungs- abschnitt 4
Adaptive Gruppe	32	29	25	20
Nicht-Adaptive Gruppe	41	50	46	12

**Tabelle 5 - Stichprobenumfang (Anzahl Tagebucheinträge) erste Forschungsfrage: Betrachtung des Nutzungszeitraums**

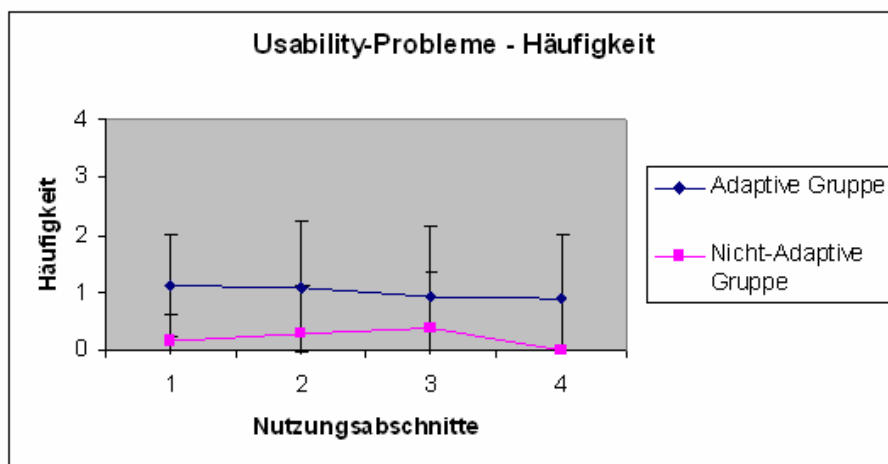
Über alle Nutzungsabschnitte des Studienzeitraums hinweg lagen die Ergebnisse der Parameter Anzahl, geschätzte Häufigkeit und Intensität von „Usability-Problemen“ bei der adaptiven Adaptiven Gruppe höher als bei der Nicht-Adaptiven Gruppe.

Abbildung 9 zeigt, dass bei der Adaptiven Gruppe zu Beginn der Studie mit einer Wahrscheinlichkeit von 0,78 „Usability-Probleme“ auftraten. Im Verlauf der Studie sank dieser Wert bis auf ein Niveau von etwa 0,5. Bei der Nicht-Adaptive Gruppe lag die Wahrscheinlichkeit über den gesamten Studienzeitraum konstant und erreichte eine höchste Wahrscheinlichkeit von 0,18.

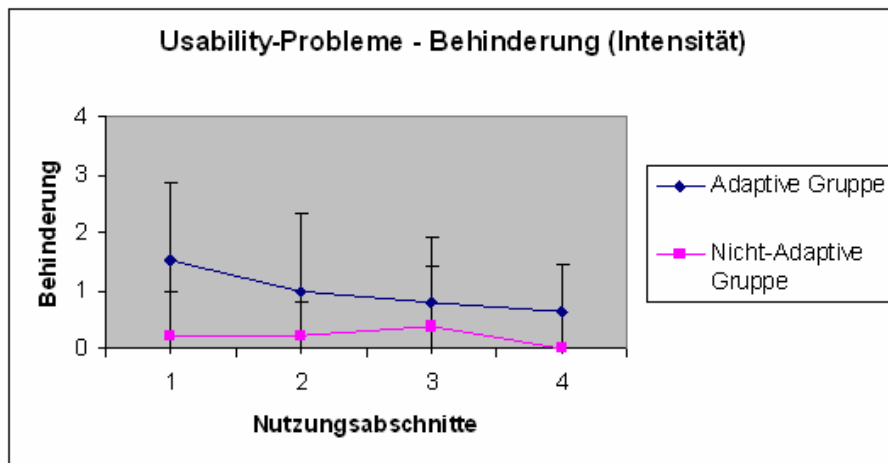


**Abbildung 9 - Erste Forschungsfrage: Ergebnis Parameter Anzahl „Usability-Probleme“ über den Studienzeitraum**

Die Mittelwerte der Parameter geschätzte Häufigkeit und Intensität wurden sowohl bei der Adaptiven als auch bei der Nicht-Adaptiven Gruppe in der unteren Hälfte der Skalen errechnet, wie die Abbildungen 10 und 11 verdeutlichen. Die Standardabweichungen lagen bei der Adaptiven Gruppe zwischen 0,9 und 1,3, bei der Nicht-Adaptiven Gruppe zwischen 0 und 0,95. Zur Beurteilung ob die Ergebnisse beider Parameter auf Zufall rückführbar oder generalisierbar waren, wurden der U-Test von Mann-Whitney durchgeführt. Es ergaben für alle Nutzungsabschnitte signifikante Werte von  $\alpha < 0,05$ .



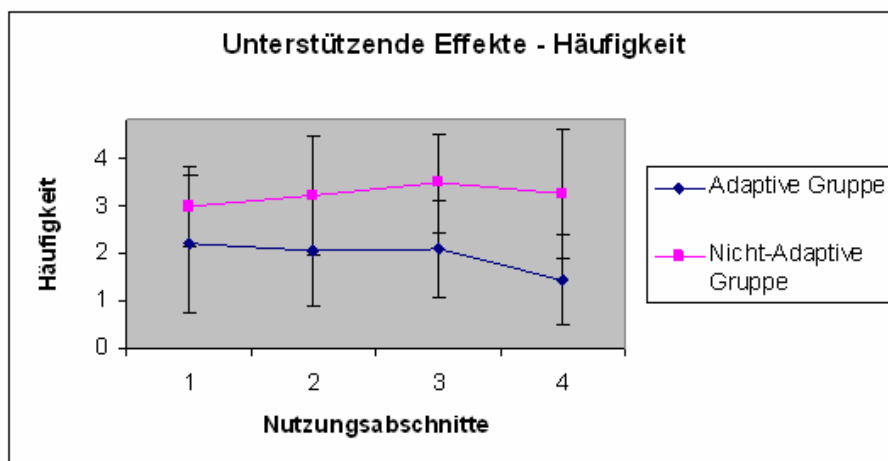
**Abbildung 10 - Erste Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Usability-Probleme“ über den Studienzeitraum**



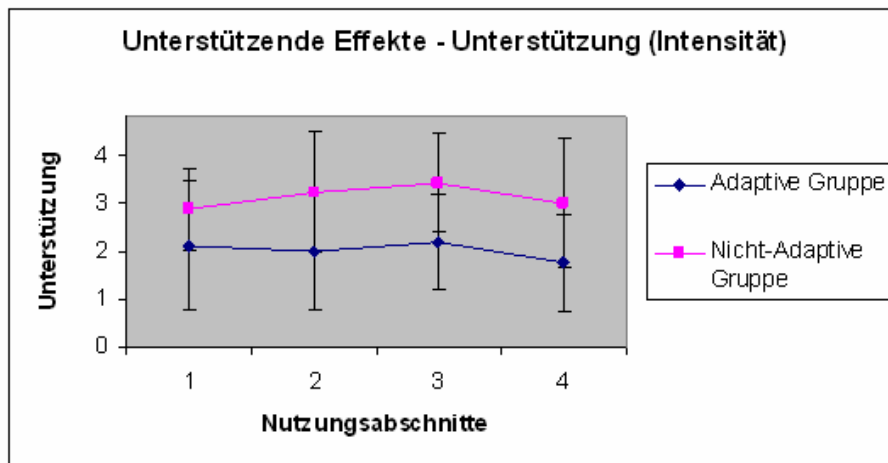
**Abbildung 11 - Erste Forschungsfrage: Ergebnis Parameter Intensität „Usability-Probleme“ über den Studienzeitraum**

Bei der Anzahl der „Unterstützenden Effekte“ befanden sich die Werte beider Testgruppen über den gesamten Testverlauf etwa im gleichen Bereich, es konnte bei fast jeder Aufgabe ein „Unterstützender Effekt“ registriert werden.

Die Abbildungen 12 und 13 weisen die errechneten Mittelwerte der Parameter geschätzte Häufigkeit und Intensität bei der Nicht-Adaptiven Gruppe jedoch um jeweils ein Drittel höher aus als bei der Adaptiven Gruppe. Die Standardabweichungen aller Werte wurden mit 0,8 bis 1,5 gemessen, für alle Vergleiche über die Nutzungsabschnitte konnten die mit dem U-Test von Mann-Whitney ermittelten Signifikanzwerte von  $\alpha < 0,05$  die deskriptiven Ergebnisse als generalisierbar bestätigen.



**Abbildung 12 - Erste Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Unterstützende Effekte“ über den Studienzeitraum**



**Abbildung 13 - Erste Forschungsfrage: Ergebnis Parameter Intensität „Unterstützende Effekte“ über den Studienzeitraum**

Um einen Verlauf der adaptiven Anpassungen innerhalb der Adaptiven Gruppe besser skizzieren zu können, wurden die Daten neben der zeitlichen Betrachtung über die Nutzungsabschnitte auch anhand der adaptiven Anpassungen der Menüs bei Microsoft Word XP analysiert.

Als Basis für die Kategorisierung der Phasen der adaptiven Anpassungen dienten die zur Dokumentation der Benutzungsoberfläche angefertigten Screenshots (vgl. Kapitel 3.3.4). Bei den Screenshots waren Verlängerungen und Verkürzungen der Menüs aufgrund adaptiver Anpassungen sowie die betroffenen Menüeinträge nachweisbar (vgl. Anhang G). Die Unterteilung erfolgte aufgrund der Beobachtung, dass nur Probanden bei denen eine wiederkehrend hohe Anzahl von mindestens 5 adaptiven Anpassungen in immerhin 2 Nutzungsabschnitten nachweisbar war, in der Abschlussbefragung angaben, sie hätten diese als Anpassungen an ihre Arbeitsgewohnheiten wahrgenommen. Dementsprechend erfolgte eine Unterteilung in vier Phasen der adaptiven Anpassung:

- Phase, in der noch keine adaptiven Anpassungen stattgefunden haben,
- Phase, in der die ersten adaptiven Anpassungen nachweisbar waren,
- Phase mit hoher Anzahl von Anpassungen (Anzahl der Anpassungen 5 und mehr),
- Phase mit einer geringen Anzahl von Anpassungen (Anzahl der Anpassungen unter 5).

Bei der Betrachtung der Phasen adaptiver Anpassungen ergab sich, aufgrund der gemachten Tagebucheinträge der Probanden in den Phasen, der in Tabelle 6 dargestellte Umfang an Stichproben für die Testgruppen:

Phasen adaptiver Anpassungen	Stichprobe
noch keine Anpassungen	14
Phase der ersten Anpassungen	35
Phase mit vielen Anpassungen	30
Phase geringer Anpassungen	27

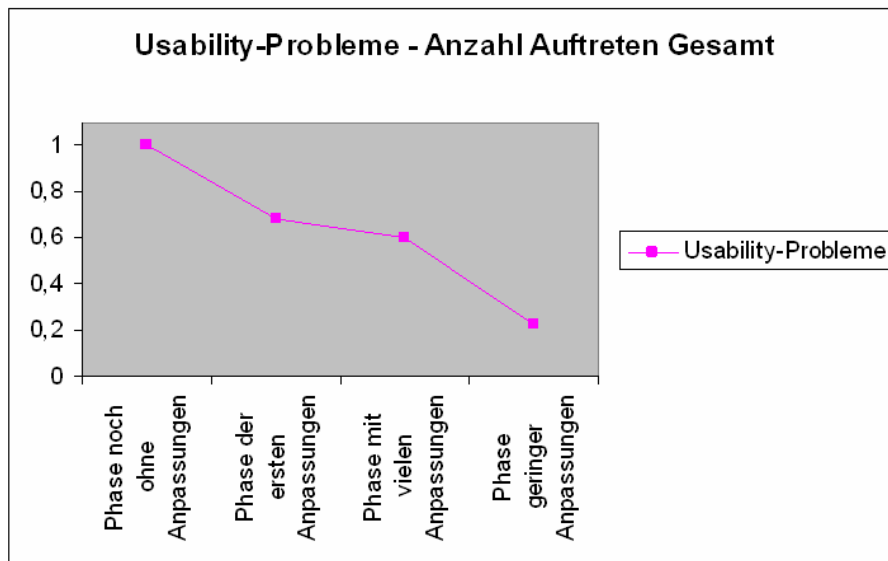
**Tabelle 6 - Stichprobenumfang (Anzahl Tagebucheinträge) erste Forschungsfrage: Betrachtung Phasen adaptiver Anpassungen bei der Adaptiven Gruppe**

Ein erneuter Blick auf die Ergebnisse der Nutzungsabschnitte zeigt bei der Betrachtung der „Usability-Probleme“ bereits eine leicht positive Entwicklung bei längerer Nutzungsdauer der Adaptivität. So traten zum Studienbeginn „Usability-Probleme“ mit einer Wahrscheinlichkeit von 0,78 pro Aufgabe auf. Dieser Wert sank während des Studienzeitraum auf 0,5 in Nutzungsabschnitt 4 bzw. 0,44 in Nutzungsabschnitt 3 (vgl. Abbildung 11). Der Parameter geschätzte Häufigkeit sank von einem Mittelwert von 1,1 zu Beginn auf 0,9 bei der letzten Datenerhebung (vgl. Abbildung 12). Ebenso zeigte der Parameter Intensität eine Tendenz zu einer geringeren Behinderung durch „Usability-Probleme“ über einen längeren Nutzungszeitraum (vgl. Abbildung 13). Der Höchststand wurde mit 1,5 in Nutzungsabschnitt 1 gemessen, im vierten und abschließenden Nutzungsabschnitt betrug der Wert 0,6. Diese Ergebnisse belegte auch ein Signifikanzwert  $\alpha < 0,05$  nach dem U-Test von Mann-Whitney für den Vergleich des ersten und vierten Nutzungsabschnitts.

Auch bei der Sichtung nach den Phasen adaptiver Anpassungen erkennt man bei den Resultaten aller Parameter Anzahl, geschätzte Häufigkeit, Intensität der „Usability-Probleme“ eine positive Entwicklung.

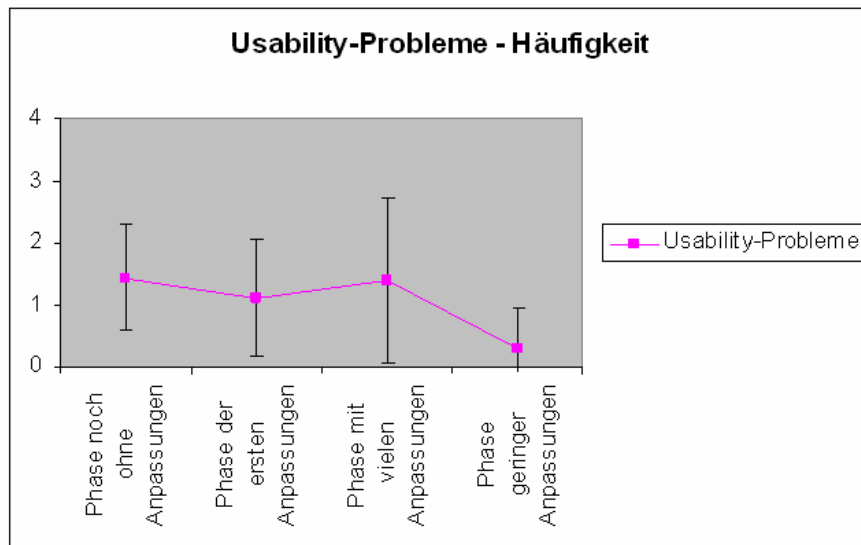
Abbildung 14 zeigt, dass in der Phase ohne Anpassungen in jeder Aufgabe ein „Usability-Problem“ auftrat, bei der Phase erster Anpassungen sowie der Phase mit vielen Anpassungen noch jeweils in 2 von 3 Aufgaben. In der Phase geringer Anpassungen wurde nur noch bei jeder fünften Aufgabe ein „Usability-Problem“ registriert. Somit näherte sich der Wert der Phase geringer Anpassungen etwa den durchschnittlichen Ergebnissen der Nicht-Adaptiven Gruppe für diesen Parameter.



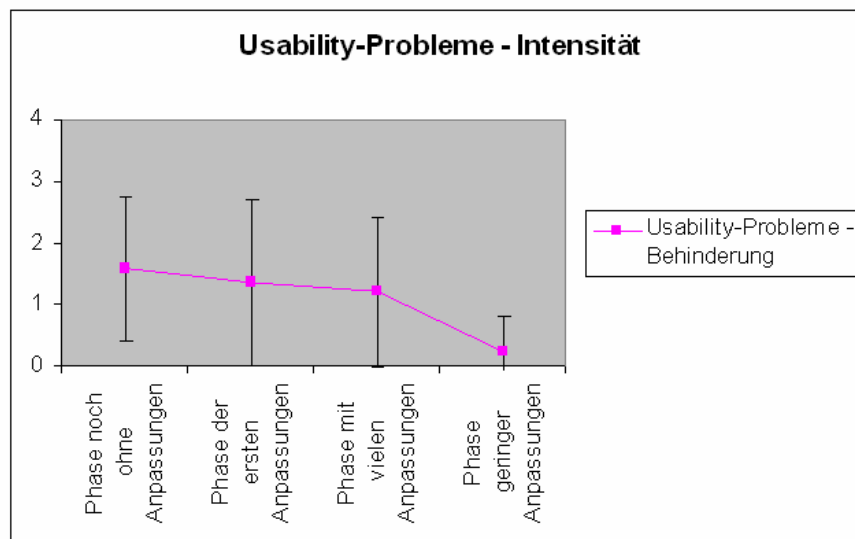


**Abbildung 14 - Erste Forschungsfrage: Ergebnis Parameter Anzahl „Usability-Probleme“ nach Phasen adaptiver Anpassung, Adaptive Gruppe**

Auch bei den Resultaten der geschätzten Häufigkeit und Intensität setzten sich diese Beobachtungen fort, wie Abbildung 15 und 16 illustrieren. So befanden sich die Mittelwerte bei beiden Parametern in der Phase ohne Anpassungen, der Phase erster Anpassungen und der Phase mit vielen Anpassungen jeweils der in mittleren Region der Skala um 1,5 mit hohen Standardabweichungen zwischen 0,8 und 1,4. Bei der Phase mit geringen Anpassungen ergaben sich für die Häufigkeit und Intensität nur noch Mittelwerte zwischen 0,2 und 0,3 mit deutlich geringeren Standardabweichung um 0,6. Diese Resultate entsprachen dem Durchschnitt der Nicht-Adaptiven Gruppe. Die Ermittlung der Signifikanz mit dem U-Test von Mann-Whitney für die deskriptiven Daten des Vergleichs der Phase ohne Anpassungen, der Phase erster Anpassungen und der Phase mit vielen Anpassungen jeweils mit der Phase geringer Anpassungen ergab immer signifikante Werte von  $\alpha < 0,05$ .



**Abbildung 15 - Erste Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Usability-Probleme“ nach Phasen adaptiver Anpassung, Adaptive Gruppe**



**Abbildung 16 - Erste Forschungsfrage: Ergebnis Parameter Intensität „Usability-Probleme“ nach Phasen adaptiver Anpassung, Adaptive Gruppe**

Eine genaue Untersuchung aufgetretener „Usability-Probleme“ bei der Adaptiven Gruppe, dargestellt in den Tabellen 7 und 8, zeigt, dass das „Behindernde Ereignis“ „Um den Menüeintrag zu finden, musste ich das Menü verlängern“ bei beiden Betrachtungsansätzen durchgängig mit der höchsten Wahrscheinlichkeit vorkam.

Studienzeitraum	Nutzungs- abschnitt1	Nutzungs- abschnitt2	Nutzungs- abschnitt3	Nutzungs- abschnitt4
„Menü verlängern“	0,47	0,47	0,44	0,45
„Menüeintrag nicht wo erwartet“	0,34	0,22	0,12	0,15
„Menü lang und unübersichtlich“	0,19	0,13	0,12	0,15
„Menüeintrag gar nicht gefunden“	0,03	0,19	0,04	0,1

**Tabelle 7 - Erste Forschungsfrage: Ergebnis Wahrscheinlichkeit einzelner „Usability-Probleme“ bei der Adaptiven Gruppen über den Studienzeitraum**

Phasen adaptiver Anpassungen	Phase ohne Anpassungen	Phase erster Anpassungen	Phase mit vielen Anpassungen	Phasen geringer Anpassungen
„Menü verlängern“	0,79	0,51	0,53	0,19
„Menüeintrag nicht wo erwartet“	0,29	0,26	0,27	0,04
„Menü lang und unübersichtlich	0,07	0,23	0,17	0,07
Menüeintrag gar nicht gefunden“	0,07	0,09	0,2	0

**Tabelle 8 - Erste Forschungsfrage: Ergebnis Wahrscheinlichkeit einzelner „Usability-Probleme“ bei der Adaptiven Gruppe nach Phasen adaptiver Anpassungen**

Dieses „Ereignis“ ist, entsprechend der Beschreibungen in Kapitel 3.4.1, auf die adaptive Funktionalität der Menüs von Microsoft Word XP zurückzuführen und konnte bei der nicht-adaptiven Systemvariante somit nicht auftreten.

Tabelle 9 veranschaulicht die Resultate der Erhebung der objektiv quantitativen Daten. Die Auswertung der ScreenCam-Aufzeichnungen ergab, dass beide Testgruppen in den Beobachtungen insgesamt etwa gleich viele Interaktionen durchgeführt hatten. Die Adaptive Gruppe benötigte für ungefähr die gleiche Anzahl an Interaktionen jedoch deutlich länger.

Zurückführen lässt sich dies auf eine größere Anzahl nicht optimaler Interaktionen, Interaktionen für die zusätzliche Arbeitsschritte benötigt wurden.

	Interaktionen Gesamt	Gesamtdauer aller Interaktionen	Nicht optimale Interaktionen (zusätzliche Arbeitsschritte benötigt)	Zusätzliche Arbeitsschritte in allen nicht optimalen Interaktionen
Adaptive Gruppe	77	00:08:23	23	124
Nicht-Adaptive Gruppe	72	00:05:29	15	49

**Tabelle 9 - Erste Forschungsfrage: Ergebnis objektiver quantitativer Daten**

Zusätzliche Arbeitsschritte ließen sich aufgrund der A/V-Aufzeichnung mit „Usability-Problemen“ erklären. Die Probanden konnten einen Menüeintrag nicht finden oder mussten das Menü aufklappen. Resultat war letztendlich die längere Gesamtdauer einer Interaktion.

### **Zweite Forschungsfrage**

In der zweiten Forschungsfrage sollte untersucht werden, ob sich Adaptivität für den Nutzer, im Hinblick auf den Umgang mit der Software, lernfördernd auswirkt.

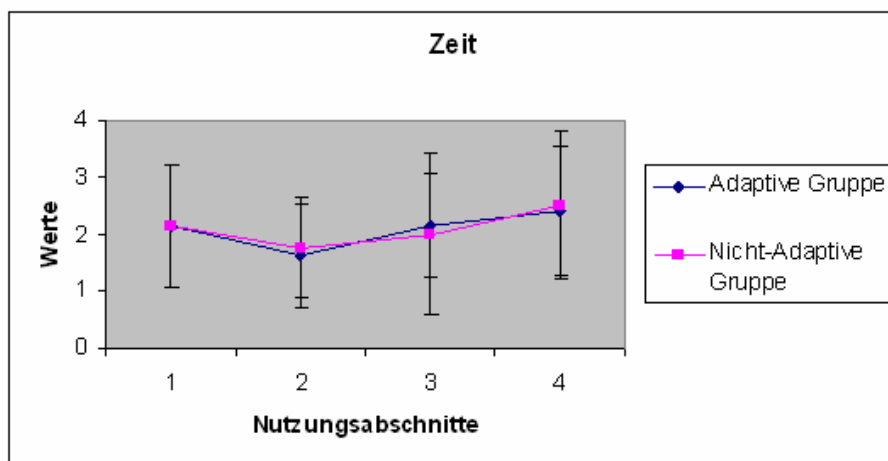
Verglichen wurden die Ergebnisse des Erhebungsinstruments Lernförderlichkeit (vgl. Kapitel 3.2.5) der Adaptiven Gruppe mit denen der Nicht-Adaptiven Gruppe über den Verlauf der Nutzungsabschnitte.

Bei der Betrachtung des Studienzeitraums ergab sich, aufgrund der durchgeführten Interventionen in den Nutzungsabschnitten, der in Tabelle 10 dargestellte Umfang an Stichproben für die Testgruppen:

Studienzeitraum	Nutzungs- abschnitt 1	Nutzungs- abschnitt 2	Nutzungs- abschnitt 3	Nutzungs- abschnitt 4
Adaptive Gruppe	7	8	7	5
Nicht-Adaptive Gruppe	7	8	7	4

**Tabelle 10 - Stichprobenumfang (Interventionen) zweite Forschungsfrage**

Die errechneten Mittelwerte beider Testgruppen waren bei den Dimensionen Zeit und Exploration annähernd gleich und lagen auf dem mittleren Level der Skalen (vgl. Abbildungen 17-20). Die Dimensionen Menge und Details und Einprägbarkeit rangierten bei der Nicht-Adaptiven Gruppe durchschnittlich leicht höher als bei der Adaptiven Gruppe.



**Abbildung 17 - Zweite Forschungsfrage: Ergebnis Lernförderlichkeit - Zeit**

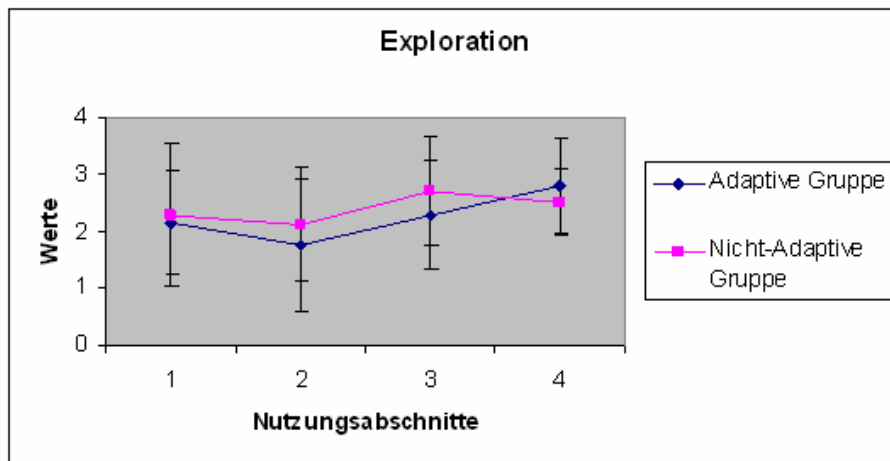


Abbildung 18 - Zweite Forschungsfrage: Ergebnis Lernförderlichkeit - Exploration

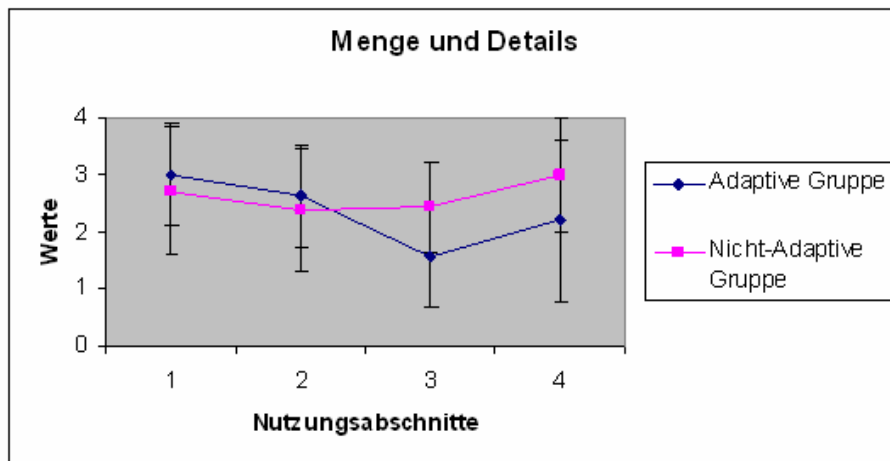
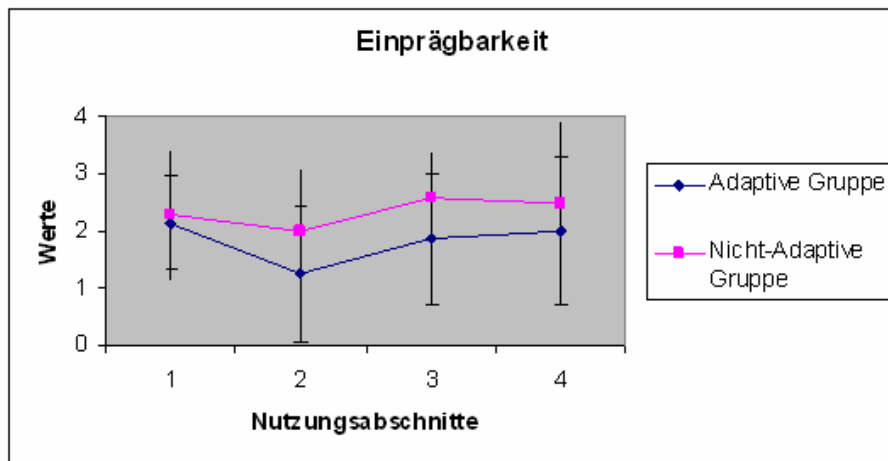


Abbildung 19 - Zweite Forschungsfrage: Ergebnis Lernförderlichkeit - Menge und Details



**Abbildung 20 - Zweite Forschungsfrage: Ergebnis Lernförderlichkeit - Einprägbarkeit**

Der Verlauf aller Dimensionen zeigte über den Studienzeitraum bei beiden Testgruppen ähnliche Tendenzen. Von der ersten Datenerhebung fielen die Werte bei allen vier Dimensionen in der zweiten bzw. dritten Nutzungsabschnitt auf den niedrigsten Stand und stiegen bis zum abschließenden Nutzungsabschnitt wieder auf das Level von Nutzungsabschnitt 1. Allerdings waren sämtliche Standardabweichungen mit Werten zwischen 0,4 und 1,5 sehr weit gestreut und teilweise recht hoch.

### **Dritte Forschungsfrage**

Die dritte Fragestellung sollte klären, ob die Probanden der Adaptiven Gruppe das adaptive System subjektiv positiver bewerten als die Nicht-Adaptive Gruppe das nicht-adaptive System.

Verglichen wurden hierfür die Ergebnisse folgender Erhebungsinstrumente:

- Einstellung gegenüber Formen der Personalisierung (vgl. Kapitel 3.2.1),
- Effizienz (vgl. Kapitel 3.2.6),
- BSMA (vgl. Kapitel 3.2.2),
- Anforderungen an adaptive Systeme (vgl. Kapitel 3.2.4),
- Anforderungen an adaptive Systeme zur finalen Einschätzung und Verbesserungspotentialen in der Abschlussbefragung (vgl. Kapitel 3.2.4).

Zur Ergebnisfindung wurden auch die erhobenen qualitativen Daten (vgl. Kapitel 3.2.8) herangezogen.

Bei der Betrachtung des Studienzeitraums ergab sich, aufgrund der gemachten Tagebucheinträge der Probanden in den Nutzungsabschnitten, der in Tabelle 11 dargestellte Umfang an Stichproben für die Testgruppen:

Studienzeitraum	Nutzungs- abschnitt1	Nutzungs- abschnitt2	Nutzungs- abschnitt3	Nutzungs- abschnitt4
Adaptive Gruppe	32	29	25	20
Nicht-Adaptive Gruppe	41	50	47	12

**Tabelle 11 - Stichprobenumfang (Tagebucheinträge) dritte Forschungsfrage: Datenerhebung Tagebuch**

Bei der Betrachtung des Studienzeitraums ergab sich, aufgrund der durchgeführten Interventionen in den Nutzungsabschnitten, der in Tabelle 12 dargestellte Umfang an Stichproben für die Testgruppen:

Studienzeitraum	Nutzungs- abschnitt1	Nutzungs- abschnitt2	Nutzungs- abschnitt3	Nutzungs- abschnitt4
Adaptive Gruppe	7	8	7	5
Nicht-Adaptive Gruppe	7	8	7	4

**Tabelle 12 - Stichprobenumfang (Interventionen) dritte Forschungsfrage: Datenerhebung Fragebogen**

Für die, im Fragebogen zum Studienabschluss (vgl. Kapitel 3.3.5.1), bewerteten Dimensionen der Anforderungen an adaptive Systeme zur Frage nach Verbesserungspotentialen des Interaktionselements ergab sich, gemäß der zu Studienbeginn vorgenommenen Aufteilung der Testgruppen, der in Tabelle 13 dargestellte Stichprobenumfang:

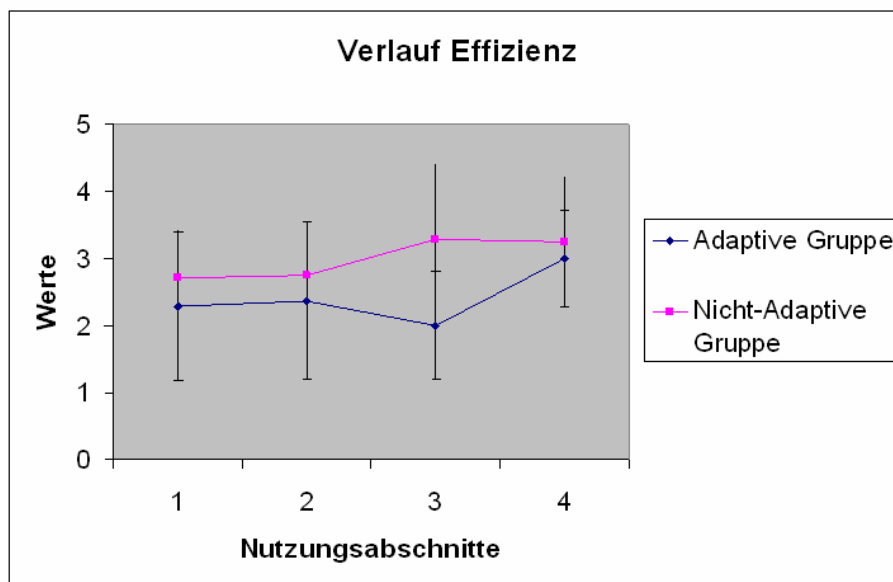
Adaptive Gruppe	8
Nicht-Adaptive Gruppe	8

**Tabelle 13 - Stichprobenumfang (Abschlussbefragungen) dritte Forschungsfrage: Datenerhebung Fragebogen zum Studienabschluss**



Die Daten der Einstellung gegenüber Formen der Personalisierung zu Studienbeginn zeigten, dass sich mit 19 von 28 eine knappe Mehrheit der Bewerber Adaptivität wünschte. Bei den ausgewählten Probanden blieben die Angaben konstant. Sowohl zu Beginn als auch beim Abschluss wünschten sich 12 Personen Adaptivität, davon jeweils immer 6 in jeder Testgruppe.

Die Werte der Effizienz lagen über den gesamten Nutzungszeitraum bei beiden Testgruppen im mittleren bis oberen Bereich der Skala, wie Abbildung 21 zeigt. Die Adaptive Gruppe verzeichnete Werte zwischen 2,2 und 3, die Nicht-Adaptive Gruppe leicht höhere Werte von 2,7 bis 3,3. Bei beiden Testgruppen konnte eine leicht positive Tendenz über den Studienverlauf verzeichnet werden. Die Standardabweichungen waren mit 0,7 bis 1,2 bei beiden Gruppen ähnlich.



**Abbildung 21 - Dritte Forschungsfrage: Ergebnis Effizienz**

Die errechneten Korrelationskoeffizienten nach Spearman für den BSMA und die Aufgabenschwierigkeit ergaben bei 6 von 15 Testpersonen Korrelationswerte von über 0,5, bei der sechzehnten Testperson reichten die erhobenen Stichproben nicht für eine Berechnung aus. Somit muss gesagt werden, dass die Bewertung der erlebten mentalen Anstrengung der Nutzung des Werkzeugs, trotz des Versuchs inhaltliche Aspekte auszuschließen, bei immerhin einem Drittel der Studienteilnehmer von der Aufgabenbearbeitung beeinflusst wurde.

Die Wertekurven des BSMA verliefen im Fragebogen und dem Tagebuch etwas unterschiedlich (vgl. Anhang K).

Im Fragebogen lagen die Ergebnisse im ersten Nutzungsabschnitt bei beiden Testgruppen noch recht deutlich auseinander. Bei der Adaptiven Gruppe war die erste Erhebung mit 67 als „etwas anstrengend“ bis „einigermaßen anstrengend“ einzustufen, der Wert der Nicht-Adaptiven Gruppe mit 29 zwischen „kaum anstrengend“ bis „etwas anstrengend“. Beide Kurven liefen im Studienverlauf „zielstrebig“ auf ihren „Treffpunkt“ bei Werten um 45 in Nutzungsabschnitt 4 zu. Somit sanken die Werte des BSMA bei der Adaptiven Gruppe, während sie bei der Nicht-Adaptiven Gruppe stetig stiegen. Beide Systeme wurden zum Studienabschluss letztendlich mit „etwas anstrengend“ bewertet.

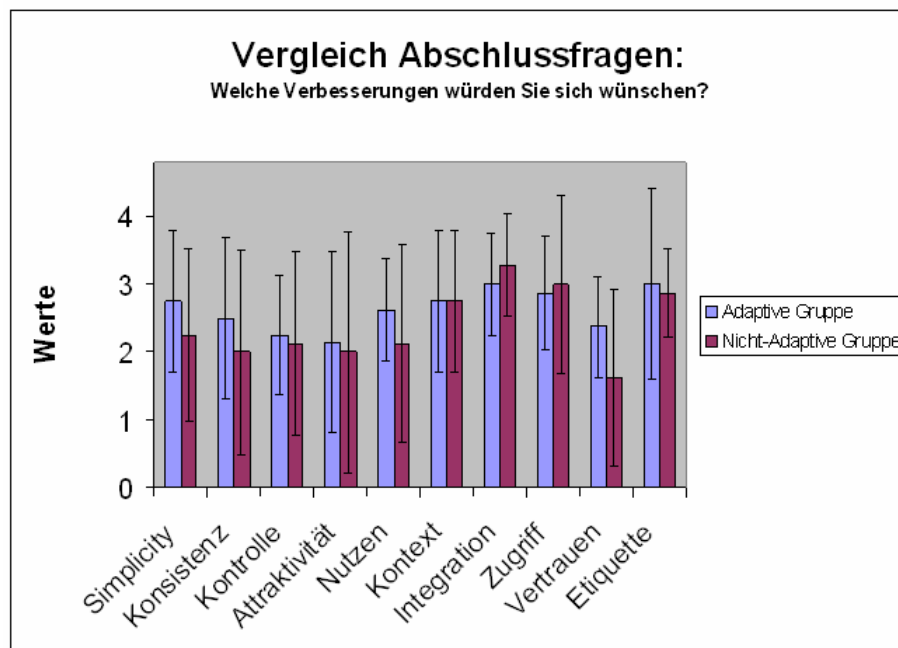
Demgegenüber blieben die Werte beider Testgruppen über den gesamten Studienverlauf im Tagebuch konstant auf einem Niveau. Die Ergebnisse bei der Adaptiven Gruppe befanden sich auf einem Level um 55, bewertet als „etwas anstrengend“ bis „einigermaßen anstrengend“. Die Bewertungen der Nicht-Adaptiven Gruppe lagen zwischen 25 und 35 bei „kaum anstrengend“ bis „etwas anstrengend“. Für die Nutzungsabschnitte 2 bis 4 konnten die Ergebnisse mit dem U-Test nach Mann-Whitney mit  $\alpha < 0,05$  als signifikant ermittelt werden.

Die Ergebnisse der Erhebungen zu den Anforderungen an adaptive Systeme lagen bei der Adaptiven Gruppe, sowohl in den Fragebogen als auch im Tagebuch über alle Nutzungsabschnitte weg, unter denen der Nicht-Adaptiven Gruppe. Alle Mittelwerte befanden sich oberhalb des Nullwertes der Skala in einem Bereich positiver Wahrnehmung (vgl. Anhang K).

In den Fragebogen ergaben sich Resultate in Bereichen zwischen 0,2 und 1,4 bei der Adaptiven Gruppe, in der Nicht-Adaptiven Gruppe zwischen 0,6 und 1,8, die Standardabweichungen rangierten zwischen 0,3 und 1,1.

Bei den Tagebuchergebnissen waren die Mittelwerte beider Testgruppen weiter voneinander entfernt als in den Fragebogen. Bei der Adaptiven Gruppe waren sie zwischen 0,3 und 1,1, bei der Nicht-Adaptiven Gruppe jeweils um mehr als 0,5 höher zwischen 1,2 und 2. Auch die Standardabweichungen der Adaptiven Gruppe schwankten mit Werten von 0,5 bis 1,4 stärker als die der Nicht-Adaptiven Gruppe mit einer maximalen Abweichung von 0,95. Bei den Dimensionen Etiquette, Privatsphäre, Kontrolle und Nutzen wurden für alle Nutzungsabschnitte signifikante Werte von  $\alpha < 0,05$  mit dem U-Test nach Mann-Whitney ermittelt, bei den Dimensionen Verständnis und Breadth of Experience jeweils für den ersten bis dritten Nutzungsabschnitt und für die Dimension Vertrauen für Nutzungsabschnitt 1 und 3.

Bei der Frage des Wunsches nach Verbesserungen im Fragebogen zum Studienabschluss waren die Werte der Adaptiven Gruppe bei den meisten abgefragten Dimensionen leicht höher als bei der Nicht-Adaptiven Gruppe, wie Abbildung 22 illustriert. Die Mittelwerte lagen bei der Adaptiven Gruppe zwischen 2,3 und 3, bei der Nicht-Adaptiven Gruppe zwischen 1,6 und 2,9, die Standardabweichungen bewegten sich zwischen 0,6 und 1,7. Ausnahmen bildeten die beiden Dimensionen Integration und Zugriff. Beide Werte lagen mit 3,2 und 3 bei der Nicht-Adaptiven Gruppe höher als bei der Adaptiven Gruppe mit 3 und 2,8.



**Abbildung 22 - Dritte Forschungsfrage: Ergebnis Abschlussfragen Anforderungen an adaptive Systeme**

Während der Studie wurde die adaptive Funktionalität von einer Testperson der Adaptiven Gruppe deaktiviert:

„Ich konnte mit den Menüs so nicht arbeiten... hat mich rasend gemacht.“ (Aussage eines Probanden aus der Adaptiven Gruppe)

Nach Beendigung der Studie deaktivierten 2 weitere Teilnehmer aus der Adaptiven Gruppe die Adaptivität. Als Gründe nannten sie, dass sie die Anpassungen als verwirrend und inkonsistent empfanden, sich den Menüs gegenüber ausgeliefert fühlten und sie durchschnittlich mehr Arbeitsschritte durch das Aufklappen der Menüs benötigten. Ein Proband äußerte sich positiv

gegenüber den wahrgenommenen Anpassungen, würde sich aber Verbesserungen in Richtung der Anforderungsdimensionen Breadth of Experience und Privatsphäre wünschen. Zwei Probanden der Adaptiven Gruppe gaben an, keine Anpassungen während der Studie wahrgenommen zu haben. Konfrontiert mit der Funktionsweise, deaktivierte einer dieser Probanden die Adaptivität. Der zweite dieser Probanden behielt das Setup bei, da er die Arbeit mit der Software während der Studie als positiv betrachtete.

Drei Probanden der Nicht-Adaptiven Gruppe entschieden sich nach der Studie die Adaptivität zu aktivieren. Einer dieser Studienteilnehmer gab an, die adaptive Funktionalität zu kennen und diese auch auf einem Privat-PC zu nutzen. Die beiden anderen Probanden kannten Adaptivität nicht und waren neugierig die Funktionalität auszuprobieren. Ein weiterer Proband der Nicht-Adaptiven Gruppe arbeitete unbeabsichtigt für 2 Nutzungsabschnitte mit angeschalteter Adaptivität. Er konnte aber keine Unterschiede zwischen beiden Varianten feststellen und gab abschließend an, dass er keine Präferenzen bezüglich der Einstellung hätte.

Zusammenfassend setzten 7 der 16 Probanden ihre Arbeit mit Microsoft Word XP nach der Studie mit aktivierter Adaptivität fort.

#### **Vierte Forschungsfrage**

Die vierte Forschungsfrage sollte klären, ob Adaptivität sich in der Einschätzung pragmatischer und hedonischer Qualität sowie Attraktivität positiv auswirkt.

Für die Datenerhebung wurde der Fragebogen AttrakDiff verwendet, welcher die Qualität eines Produkts in den vier Dimensionen Pragmatische Qualität (PQ), Hedonische Qualität Identität (HQ-I), Hedonische Qualität Stimulation (HQ-S) und Attraktivität (ATT) abfragt (Hassenzahl, Burmester, Koller, 2003).

Bei der Betrachtung des Studienzeitraums ergab sich, aufgrund der durchgeführten Interventionen in den Nutzungsabschnitten, der in Tabelle 14 dargestellte Umfang an Stichproben für die Testgruppen:

Studienzeitraum	Nutzungs- abschnitt1	Nutzungs- abschnitt2	Nutzungs- abschnitt3	Nutzungs- abschnitt4
Adaptive Gruppe	7	8	7	5
Nicht-Adaptive Gruppe	7	8	7	4

**Tabelle 14 - Stichprobenumfang (Interventionen) vierte Forschungsfrage**

Alle vier Dimensionen hatten sowohl bei der Adaptiven Gruppe als auch bei der Nicht-Adaptiven Gruppe im ersten Nutzungsabschnitt vergleichbare Mittelwerte im neutralen Bereich der Skala. Wie Tabelle 15 veranschaulicht, zeichnete sich zum Studienende bei allen vier Dimensionen in beiden Testgruppen ein Trend zu einer positiveren Wahrnehmung ab, der bei der Nicht-Adaptiven Gruppe jedoch deutlich stärker ausfiel als bei der Adaptiven Gruppe:

Dimension AttrakDiff	Testgruppe	Nutzungs- abschnitt1	Nutzungs- abschnitt4
PQ	Adaptive Gruppe	0,16	0,71
	Nicht-Adaptive Gruppe	0,24	1,42
HQ-I	Adaptive Gruppe	0,51	0,51
	Nicht-Adaptive Gruppe	0,34	0,85
HQ-S	Adaptive Gruppe	-0,14	0,11
	Nicht-Adaptive Gruppe	-0,02	1,07
ATT	Adaptive Gruppe	0,63	0,88
	Nicht-Adaptive Gruppe	0,46	1,35

**Tabelle 15 - Vierte Forschungsfrage: Vergleich der Dimensionen des AttrakDiff**

### **Fünfte Forschungsfrage**

Die Studie zur Adaptivität bei Interaktionselementen von Burmester (Burmester, 2003) ergab, dass keine signifikanten Unterschiede für die Nutzung von adaptiver Software zwischen Nutzergruppen mit unterschiedlichen Nutzungserfahrungen bestehen. Diese Ergebnisse sollten in der fünften Fragestellung überprüft werden.

Es erfolgte ein Vergleich der Nutzergruppen Experten, User, Novizen (vgl. Kapitel 3.4.2) innerhalb der Adaptiven Gruppe. Zur Beantwortung herangezogen wurden die Ergebnisse der folgenden Datenerhebungen:

- Anzahl, geschätzte Häufigkeit und Intensität von „Usability-Problemen“ und „Unterstützenden Effekten“ (vgl. Kapitel 3.2.3),
- Anforderungen an adaptive Systeme (vgl. Kapitel 3.2.4),
- BSMA (vgl. Kapitel 3.2.2),
- Effizienz (vgl. Kapitel 3.2.6.),
- Lernförderlichkeit (vgl. Kapitel 3.2.5).

Außerdem wurden die erhobenen qualitativen Daten (vgl. Kapitel 3.2.8) untersucht.

Bei der Betrachtung des Studienzeitraums ergab sich, aufgrund der gemachten Tagebucheinträge der Probanden in den Nutzungsabschnitten, der in Tabelle 16 dargestellte Umfang an Stichproben für die Nutzergruppen:

Studienzeitraum	Nutzungs- abschnitt1	Nutzungs- abschnitt2	Nutzungs- abschnitt3	Nutzungs- abschnitt4
Experten	7	8	9	5
User	19	10	6	4
Novizen	6	11	10	11

**Tabelle 16 - Stichprobenumfang (Tagebucheinträge) fünfte Forschungsfrage: Datenerhebung Tagebuch**

Bei der Betrachtung des Studienzeitraums ergab sich, aufgrund der durchgeführten Interventionen in den Nutzungsabschnitten, der in Tabelle 17 dargestellte Umfang an Stichproben für die Nutzergruppen:

Studienzeitraum	Nutzungs- abschnitt1	Nutzungs- abschnitt2	Nutzungs- abschnitt3	Nutzungs- abschnitt4
Experten	2	3	3	1
User	3	3	2	2
Novizen	2	2	2	2

**Tabelle 17 - Stichprobenumfang (Interventionen) fünfte Forschungsfrage: Datenerhebung Fragebogen**

Der Vergleich der Parameter von „Unterstützenden Effekten“ zeigt über den Studienverlauf eine Tendenz zu Gunsten der Experten. Lagen sämtliche Parameter im ersten Nutzungsabschnitt noch unter denen der User und der Novizen, stiegen die Werte im Verlauf der Studie über bzw. auf das Niveau der Vergleichsdaten der beiden anderen Nutzergruppen.

Bei Usern und Novizen rangierten die Parameter Anzahl und Intensität auf ähnlichen Levels, bei der geschätzten Häufigkeit lagen die Werte der Novizen über denen der User. Mittels des Mediantests für k unabhängige Stichproben wurden signifikante Werte von  $\alpha < 0,05$  in den Nutzungsabschnitten 1 bis 3 des Parameters geschätzte Häufigkeit und in den Nutzungsabschnitten 1 und 4 des Parameters Intensität errechnet.

Bei den „Usability-Problemen“ befanden sich die Ergebnisse der Parameter Anzahl, geschätzte Häufigkeit und Intensität bei den Experten über denen der User, deren Ergebnisse wiederum über denen der Novizen. Durch den Mediantest für k unabhängige Stichproben ergaben sich signifikante Werte von  $\alpha < 0,05$  für die Ergebnisse in den Nutzungsabschnitten 2 bis 4 des Parameters Häufigkeit und in alle Nutzungsabschnitten des Parameters Intensität (vgl. Anhang L).

Die Dimensionen der Anforderungen an adaptive Systeme wurden, sowohl im Tagebuch als auch in den Fragebogen, von den Experten in den meisten Fällen negativer bewertet als von Usern und Novizen, die Werte dieser beiden Gruppen sind in vielen Fällen annähernd gleich. Signifikante Werte von  $\alpha < 0,05$  durch den Mediantest für k unabhängige Stichproben für die Ergebnisse aus dem Tagebuch konnten jedoch nur in Einzelfällen errechnet werden. Besonders deutlich wurden unterschiedliche Bewertungen bei der Dimension Breadth of Experience, hier ergaben sich signifikante Werte für die Nutzungsabschnitte 1 bis 3. Eine Ausnahme bildete die

Dimension Privatsphäre, die Werte ließen hier keine klare Tendenz zu, es konnte auch in keinem Nutzungsabschnitt ein signifikanter Wert errechnet werden (vgl. Anhang L).

Klare Unterschiede in der Bewertung konnten beim BSMA erkannt werden. Die errechneten Mittelwerte der Experten lagen zwischen 155 zu Beginn und 60 im letzten Studienabschnitt bei „einigermaßen anstrengend“ bis „stark anstrengend“. Bei den Usern lagen die Mittelwerte zwischen 30 und 70 bei „etwas anstrengend“ bis „einigermaßen anstrengend“. Der ermittelte Höchstwert bei den Novizen lag bei 43 und somit bei „etwas anstrengend“ (vgl. Anhang L). Für die Ergebnisse im Tagebuch konnte für die Nutzungsabschnitte 3 und 4 ein Signifikanzniveau von  $\alpha < 0,05$  mit dem Mediantest für k unabhängige Stichproben errechnet werden.

Bei der Effizienz lagen die Mittelwerte der Experten während der gesamten Studiendauer unter denen der User und Novizen. Die Mittelwerte der User lagen in den ersten beiden Nutzungsabschnitten unter denen der Novizen, stiegen im dritten und vierten Abschnitt jedoch auf ein höheres Niveau.

Bei den Dimensionen Zeit, Menge und Details und Einprägbarkeit der Lernförderlichkeit lagen die Ergebnisse der Experten unter denen der beiden anderen Nutzergruppen. Der Vergleich zwischen Usern und Nutzern zeigte für die Dimensionen Zeit und Einprägbarkeit keine klaren Unterschiede, bei der Dimension Menge und Details lagen die Werte der User über denen der Novizen. Die Ergebnisse der Dimension Exploration zeigten generell keine klaren Tendenzen.

Die Untersuchung der qualitativen Daten ergab, dass es sich bei 2 der 4 Probanden, die das adaptive Setup deaktivierten, um Experten handelte. Die beiden anderen Teilnehmer waren User.

Ein weiterer User bewertete das adaptive System positiv, sah aber auch Verbesserungspotentiale.

Jeweils ein Novize und ein User registrierten die adaptiven Anpassungen nicht. Der Novize äußerte sich abschließend jedoch positiv über das System, der User gehörte zu den Probanden, die Adaptivität deaktivierten.



## Ergebnisse der qualitativen Datenerhebung zur Evaluation der Erhebungsinstrumente

Generell wurden der Studienablauf und die Erhebungsinstrumente von den Probanden verstanden und recht positiv aufgenommen. Einzig bei dem Erhebungsinstrument „Usability-Probleme“ und „Unterstützende Effekte“ ergab sich eine Tendenz zu eher negativen Einschätzungen und Beobachtungen mit unterschiedlich starken Auswirkungen auf die Datenerhebung.

Unterschiede zwischen einer der Test- oder Nutzergruppen konnten nicht festgestellt werden. Die folgende Tabelle 18 präsentiert die Ergebnisse der in Kapitel 3.6 beschriebenen Kategorien.

Kategorie	Beschreibung qualitativer Daten
Allgemein	Insgesamt bewerteten die Probanden den Studienablauf und die Betreuung als positiv, sie hatten wenig Probleme mit dem Ausfüllen der Formulare.
Bewertung im Mittel in den Fragebogen zur Intervention	Zwei Probanden empfanden die Bewertung der Anforderungen an adaptive Systeme im Mittel für die „Ereignisse“ als schwierig. Der Bezug zur Bewertung der „Ereignisse“ auf alle bisherigen Erfahrungen war ihnen unklar. Die Fragestellung wurde eher auf die letzte Aufgabe bezogen und als missverständlich formuliert empfunden. Die übrigen 14 Probanden hatten keine Probleme mit der Fragestellung.
Erhebungsinstrument Anforderungen an adaptive Systeme	Ein Proband empfand die Fragestellung als generell verwirrend. Ein weiterer Proband konnte die „eher menschlichen Attribute“ nur schwer auf das Werkzeug PC anwenden. Zwei Probanden verstanden die Fragestellung der Dimension Breadth of Experience nicht bzw. gaben an diese als nicht zutreffend zu beurteilen. Die Dimension Privatsphäre betrachtet ein Proband als unpassend, ein anderer gab an, dass diese erst durch die Fragestellung in sein Bewusstsein gerückt sei.
Erhebungsinstrument AttrakDiff	Ein Proband konnte die „eher menschlichen Attribute“ nur schwer auf das Werkzeug PC anwenden. Es handelt sich dabei um den selben Proband, der dies auch gegenüber dem Erhebungsinstrument Anforderungen an adaptive Systeme äußerte. Ein weiterer Proband empfand einige Wortpaare als unpassend Es

	handelte sich um die Paare: fachmännisch-laienhaft, hässlich-schön, mutig-vorsichtig und harmlos-herausfordernd.
Erhebungsinstrument „Usability-Probleme“ und „Unterstützende Effekte“	<p>Die Ereignisbeschreibung "Kurz und übersichtlich" wurde von einem Proband als nicht immer zutreffend empfunden, da „... die Übersichtlichkeit auch durch die Trennstriche bei den Menüs hergestellt wird, selbst wenn diese lang sind“ (Aussage eines Probanden).</p> <p>Das „Unterstützendes Ereignis“ hätte ein weiterer Teilnehmer anders formuliert. Anstatt "Wo erwartet" hätte der Proband das „Ereignis“ als "Ich wusste was ich will" beschrieben.</p> <p>Von einem Probanden wurde die Ereignisbewertung zwar verstanden, aber nicht direkt mit den Veränderungen des Interaktionselements in Zusammenhang gebracht.</p> <p>Vier Probanden hatten Probleme mit dem Verständnis, dass für die Bewertung der „Ereignisse“ nur die Menüs und nicht auch die Symbole berücksichtigt wurden. Eine dieser Testpersonen äußerte den Wunsch nach einem Textfeld für eigene Beschreibungen der „Ereignisse“.</p> <p>Die Auswertung der Kategorie Wahrnehmung der qualitativen Daten ergab außerdem, dass 4 Probanden der Adaptiven Gruppe und 2 Probanden der Nicht-Adaptiven Gruppe „Ereignisse“ während der Vor-Ort-Beobachtung teilweise abweichend von den objektiven Einschätzungen des Studienleiters wahrgenommen hatten.</p> <p>Bei 2 Probanden wurde festgestellt, dass durch lange Dauer und viele Interaktionen bei einer Aufgabenbearbeitung, „Ereignisse“ letztendlich nur schwer zuzuordnen waren.</p> <p>Insgesamt ergaben sich für 10 Probanden Beeinträchtigungen unterschiedlicher Auswirkungen bei der Datenerhebung mit diesem Instrument.</p>
Führung des Tagebuchs	<p>Ein Proband beurteilte die Fragen als monoton.</p> <p>Das Online-Formular des Tagebuchs wurde allgemein als sehr praktisch empfunden.</p>

**Tabelle 18 - Ergebnisse zur Evaluation der Erhebungsinstrumente**

## 5 Diskussion

### *Schlechtere subjektive Bewertung der Adaptivität gegenüber dem nicht-adaptiven System*

Zusammenfassend zeigen die Ergebnisse der Studie, dass das adaptive System bei der überwiegenden Zahl der Erhebungsinstrumente schlechter bewertet wurde als das nicht-adaptive System. Viele dieser Resultate wurden auch durch die Berechnung von Signifikanzwerten zusätzlich belegt. So traten Usability-Probleme in größerer Zahl, häufiger und mit stärkeren Auswirkungen auf, während unterstützende Effekte gleichzeitig seltener und mit geringeren positiven Effekten die Arbeit zu unterstützen vermochten. Die Ergebnisse ergaben auch, dass sich Adaptivität nicht lernfördernd auf die Nutzung des Systems auswirkt. Die Probanden gaben sogar an, dass eine größere Menge an Details der Software erlernt werden müsse und Gelerntes sich schlechter einprägt. Von den Probanden wurde das adaptive System gegenüber der nicht-adaptiven Variante als weniger effizient und anstrengender in der Handhabung bewertet. Auch die, für die Gebrauchstauglichkeit von adaptiven Systemen, wichtigen Aspekte Nutzen, Etikette, Privatsphäre, Verständnis, Vertrauen, Breadth of Experience und Kontrolle wiesen negativere subjektive Einschätzungen auf. Weiter deuteten die Ergebnisse des AttrakDiff Adaptivität als allgemein weniger gebrauchstauglich und zufriedenstellend, die Probanden identifizierten sich weniger mit dem Produkt und fühlten sich auch geringer stimuliert.

Diese Ergebnisse wurden auch durch die stichprobenartige Erhebung objektiv quantitativer Daten zu Interaktionsdauer und Bearbeitungsschritten untermauert.

Vier Probanden der Adaptiven Gruppe deaktivierten die Adaptivität nach Studienabschluss und nur 7 der 16 Studienteilnehmer wollten mit aktivierter Adaptivität weiterarbeiten. Davon waren 2 Probanden aus der Nicht-Adaptiven Gruppe, die eine adaptive Funktionalität bei dem Testprodukt nicht kannten und nur aus Neugier umstellten.

### *Wunsch nach Adaptivität*

Die Neugier dieser beiden Probanden zeigt aber auch ein generelles Interesse von Nutzern an Adaptivität. Zu vergleichbaren Ergebnissen kamen auch die Studien von Dubiella (Dubiella, 2003) und von McGrenere, Findlater (McGrenere, Findlater, 2004). So äußerten auch 19 von 28 Studienbewerber den Wunsch nach einem adaptiven System. Von den 16 Probanden wünschten sich sowohl vor als auch nach der Studie 12 Personen adaptive Systeme.

### *Abweichungen zwischen Nutzergruppen mit unterschiedlichen Nutzungserfahrungen*

Die Ergebnisse der Studie zur Adaptivität bei Interaktionselementen von Burmester (Burmester, 2003) zeigten keine signifikanten Unterschiede zwischen Nutzergruppen mit verschiedenen Nutzungserfahrungen.

Die Überprüfung dieser Ergebnisse, anhand der Erhebungsinstrumente „Usability-Probleme“ und „Unterstützende Effekte“, Anforderungen an adaptive Systeme, BSMA, Effizienz und Lernförderlichkeit, tendierten jedoch zu subjektiv schlechteren Bewertungen bei Probanden mit zunehmend umfangreicheren Nutzungserfahrungen. Probanden mit einem hohen Kenntnisstand schalteten die adaptive Funktionalität nach Studienabschluss auch eher ab als Nutzer mit geringeren Erfahrungen. Die sogenannten Novizen, Nutzer mit geringeren Nutzungskenntnissen, nahmen adaptive Anpassungen allerdings auch seltener wahr als die anderen Nutzergruppen. Beobachtungen, die auch bereits McGrenere, Baecker und Booth (McGrenere, Baecker, Booth, 2002) gemacht hatten.

### *Einstiegshürden für die Nutzung adaptiver Systeme*

Neben dem Problem, dass Nutzer Adaptivität nicht wahrnehmen und somit auch nicht effektiv nutzen können, belegten die Ergebnisse aber auch das, von Burmester (Burmester, 2003) identifizierte, Problem bestehender Einstiegshürden. Der skizzierte Verlauf von „Usability-Problemen“ über den Studienzeitraum und anhand der Phasen adaptiver Anpassungen verdeutlicht, dass Adaptivität gerade in den ersten Nutzungsabschnitten und Phasen als besonders behindernd wahrgenommen wird. Will ein Hersteller durch Adaptivität die Loyalität seiner Kunden erhöhen (Riecken, 2000), müssen Wege gefunden werden, zu allererst diese Hürden aus dem Weg zu räumen.

### *Notwendigkeit adaptive Funktionalität mehr an die Vorstellungen der Nutzer anzupassen*

Eine Notwendigkeit die Gestaltung von adaptiven Systemen und adaptiver Funktionalität zu verbessern, unterstreichen generell sämtliche Ergebnisse der Studie, speziell auch die recht hohen Werte zu den Verbesserungspotentialen der Anforderungen an adaptive Systeme in der Abschlussbefragung.

Wer Adaptivität sinnvoll umsetzen will, sollte mehr auf die tatsächlichen Wünsche und Bedürfnisse der Nutzer eingehen. Dies betrifft zum Einen die allgemeine Umsetzung entsprechend den Anforderungen an adaptive Systeme (vgl. Kapitel 2.2.3). Weiter sollte den Nutzern aber auch mehr Möglichkeiten gegeben werden auf die Adaptivität einwirken zu

können, wie es Horvitz beispielsweise mit seiner Idee für Mixed-Initiative UIs liefert (Horvitz, 1999).

#### *Notwendigkeit die gestalterische Umsetzung der Adaptivität zu optimieren*

Die Ergebnisse der Dimension HQ-S des AttrakDiff verdeutlichen, dass die adaptiven Menüs von Microsoft Word XP die Probanden weniger stimulierten und zur Exploration anregten als die statischen Menüs der nicht-adaptiven Variante. Die Vermutung liegt nahe, dass dieses Ergebnis auf die Verkürzung der Menüs durch die adaptive Funktionalität zurückzuführen ist. Nicht verwendete Menüeinträge werden dabei in einen nicht sichtbaren Abschnitt des Menüs verschoben. Befindet sich ein gesuchter Menüeintrag nicht im unmittelbar sichtbaren Teil des Menüs, müssen diese manuell oder automatisch vom System aufgeklappt werden.

Es zeigte sich, dass die Mehrzahl der aufgetretenen „Usability-Probleme“ bei der Adaptiven Gruppe unmittelbar auf diese gestalterische Umsetzung zurückzuführen waren. So basierte die Mehrheit der aufgetretenen „Usability-Probleme“ aus dem definierten „Ereignis“ „Um den Menüeintrag zu finden, musste ich das Menü verlängern“. Außerdem gab auch ein Proband der Adaptiven Gruppe explizit an, die Adaptivität aus diesem Grund nach Studienende deaktivieren zu wollen.

„... da ich die Anpassungen der Menüs als störend empfinde und mehr Arbeitsschritte zum Ausführen einer bekannten Funktion benötige.“ (Aussage eines Probanden der Adaptiven Gruppe)

So besteht die Notwendigkeit die gestalterische Umsetzung von Adaptivität dahingehend zu optimieren, dass Adaptivität nicht für „neue“ Usability-Probleme verantwortlich ist und gleichzeitig die Stimulation zur Exploration der Software vermindert.

## 6 Ausblick

Die Ergebnisse zeigen insgesamt den Wunsch der Nutzer nach adaptiven Systemen. Demgegenüber stehen jedoch die schlechteren Ergebnisse des adaptiven Testprodukts verglichen mit der nicht-adaptiven Variante. Die Betrachtung der Bewertungen des adaptiven Systems zeigen, dass die negativen Resultate des adaptiven Testprodukts zu einem großen Teil mit gestalterischen Mängeln der Umsetzung zusammenhängen. So basiert die Mehrheit aufgetretener Usability-Probleme auf der Gestaltung des Interaktionselements. Auch Einstiegshürden für die Nutzung und das Verhindern der Exploration der Software können auf die Gestaltung zurückgeführt werden.

Ein Lösungsansatz, um diese Probleme zu minimieren, ist die Erstellung von Usability-Richtlinien für die Gestaltung adaptiver Systeme. Durch die Erstellung spezifischer Erhebungsinstrumente zur Evaluation adaptiver Systeme hat die Studie hier einen ersten Schritt getan. Das Feedback der Probanden zeigt, dass die Fragestellungen mehrheitlich verstanden wurden. Für weitere Studien sollten die, im Studienverlauf geäußerten, Verbesserungsvorschläge der Probanden und die Ergebnisse des Seminars zur Senkung von Einstiegshürden für die Nutzung adaptiver Systeme der Hochschule der Medien in Stuttgart (Burmester, Wissmann, 2004) hinzugezogen werden. Das Ergebnis sollte in einer weiteren Studie evaluiert werden.

Bei einer weiterführenden Studie sollten auch Log-Dateien oder spezifische Softwaretools eingesetzt werden, die Nutzungszeiten und Bearbeitungsschritte über die gesamte Studiendauer registrieren, wie z.B. in den Untersuchung von McGrenere, Baecker und Booth (McGrenere, Baecker, Booth, 2002) und Höök (Höök, 1997). So lassen sich vor allem Usability-Probleme genauer bestimmen. Die Datenerhebung auf der Basis subjektiver Einschätzungen wies in der Studie teilweise Defizite bezüglich dem Verständnis von Fragestellungen und der Wahrnehmung von definierten Usability-Problemen seitens der Probanden auf. Auch wenn die stichprobenartig durchgeführten Erhebungen objektiver quantitativer Daten in den Vor-Ort-Beobachtungen letztlich die subjektiven Einschätzungen der Erhebungsinstrumente zu den Usability-Problemen untermauerten.

Die Studie verdeutlichte, dass sich unterschiedliche Nutzungserfahrungen auch auf die Wahrnehmung und Bewertung adaptiver Systeme auswirken können. Die Ergebnisse von Dubiella (Dubiella, 2003) zeigten ebenfalls, dass verschiedene Zielgruppen unterschiedliche Anforderungen an ein Produkt stellen. Dies entspricht dem Grundprinzip der Personalisierung. Die Studienergebnisse weisen jedoch daraufhin, dass dieser Aspekt bei dem Testprodukt nicht

genügend berücksichtigt wurde. So sollten im Usability-Engineering Prozess bereits bei der Anforderungsanalyse Fragen zu möglichen adaptiven Features geklärt werden. Dies betrifft z.B. den Wunsch möglicher Nutzergruppen nach adaptiven Funktionalitäten, ob Adaptivität die Gebrauchstauglichkeit eines Produkts wirklich verbessert, wo sie eingesetzt und wie gestalterisch umgesetzt werden kann.

## **7 Zusammenfassung**

Personalisierbare Produkte erlauben es dem Anwender das ganze Produkt oder einzelne Elemente seinen Wünschen und Arbeitsweisen entsprechend anzupassen.

Innerhalb der Personalisierung beschreibt die Adaptivität eine automatisierte Form der Personalisierung, in der das System initiativ ein Modell des Nutzers aufbaut und sich selbständig den Gewohnheiten des Nutzers anpasst.

In der vorliegenden Arbeit sollten ein Konzept und Erhebungsinstrumente zur Untersuchung adaptiver UI-Elemente entwickelt werden. Erhoben wurden dabei in erster Linie subjektive Einschätzungen der Probanden, ergänzt durch qualitative Daten aus Kommentaren der Testpersonen und Beobachtungen des Testleiters sowie stichprobenartig objektive quantitative Daten. Bei der Erhebung der subjektiven Einschätzungen zum Einsatz kamen die Standardfragebogen AttrakDiff und BSMA sowie speziell entwickelte Fragebogen zu Anforderungen an adaptive Systeme, zur Einstellung gegenüber Formen der Personalisierung, zur Lernförderlichkeit, Effizienz und Anzahl, Häufigkeit und Intensität von Usability-Problemen und unterstützenden Effekten.

Untersucht werden sollten die Auswirkungen von Adaptivität bei Interaktionselementen über einen längeren konstanten Gebrauch eines Produkts in einer Feldstudie. Als Testprodukt wurden die Menüs von Microsoft Word XP verwendet. Das Konzept sah in erster Linie die Führung eines Tagebuchs durch die Testpersonen vor, ergänzt durch periodische Vor-Ort-Interventionen am Arbeitsplatz. Die Interventionen unterteilen sich in einen Fragebogen-Interview-Teil und eine Beobachtung der Probanden bei der Aufgabenbearbeitung. Die Untersuchung wurde durch eine Abschlussbefragung beendet.

Die Studie hatte einen weitgehend deskriptiv-explorativen Charakter und sollte der Verbesserung von Verfahren zur Evaluation von adaptiven Systemen dienen. Ein weiteres Ziel bestand aber auch in der Überprüfung von Ergebnissen anderer Forschungsarbeiten sowie der Generierung erster relevanter Erkenntnisse zu den aufgestellten Forschungsfragen. Unterschiede zwischen adaptiven und nicht-adaptiven Systemen sollten durch die Aufteilung der Probanden in eine Adaptive Gruppe, die mit einem adaptiven System arbeitete, und eine Nicht-Adaptive Gruppe, die mit einem identischen nicht-adaptiven System arbeitete, dargestellt werden.



Die erste Fragestellung verfolgte das Ziel mögliche, durch die Adaptivität auftretende Usability-Probleme entsprechend ihrer Art, Häufigkeit und Auswirkungen zu identifizieren und mögliche Erklärungen für das Auftreten zu lokalisieren. Ein Vergleich der beiden Testgruppen ergab, dass durch die Adaptivität Usability-Probleme in größerer Zahl, häufiger und mit stärkeren Auswirkungen auftraten, während unterstützende Effekte gleichzeitig seltener und mit geringeren positiven Effekten die Arbeit zu unterstützen vermochten. Der skizzierte Verlauf von Usability-Problemen über den Nutzungszeitraum und anhand des Verlaufs adaptiver Anpassungen verdeutlichte, dass Adaptivität gerade in den ersten Nutzungsabschnitten als besonders hindernd wahrgenommen wurde. Es zeigte sich auch, dass die Mehrzahl der aufgetretenen Usability-Probleme unmittelbar auf die gestalterische Umsetzung der adaptiven Funktionalität der Menüs von Microsoft Word XP zurückzuführen waren.

Die zweite Fragestellung sollte klären, ob adaptive UI-Elemente sich für die Nutzung eines Systems lernfördernd auswirken. Die Probanden mit angeschalteter Adaptivität hatten jedoch den Eindruck, dass eine größere Menge an Details der Software erlernt werden musste und sich Gelerntes schlechter einprägte.

Die dritte Fragestellung sollte evaluieren ob Probanden das adaptive System subjektiv positiver bewerten als das System ohne Adaptivität und sich Adaptivität wünschen. Von den Probanden wurde das adaptive Systeme gegenüber der nicht-adaptiven Variante als weniger effizient und anstrengender in der Handhabung bewertet. Auch die für die Gebrauchstauglichkeit von adaptiven Systemen wichtigen Aspekte Nutzen, Etikette, Privatsphäre, Verständnis, Vertrauen, Breadth of Experience und Kontrolle wiesen negativere subjektive Einschätzungen auf. Allerdings wünschten sich 19 von 28 Studienbewerber ein adaptives System. Auch von den Probanden wünschten sich sowohl vor als nach der Studie 12 von 16 Personen Adaptivität.

Weiter sollte die Frage geklärt werden, ob adaptive Funktionen sich auf die Wahrnehmung von Attraktivität und hedonischer Qualitäten von Seiten der Nutzer positiv auswirken. Hier zeichnete sich bei beiden Testgruppen ein positiver Trend über die Studiendauer ab, der bei dem nicht-adaptiven System deutlich stärker ausfiel als bei dem adaptiven.

Die letzte Fragestellung untersuchte, ob es Unterschiede in der Bewertung von Adaptivität bei Nutzergruppen mit verschiedenen Nutzungserfahrungen gibt. Die Ergebnisse zeigten eine Tendenz zu subjektiv schlechteren Bewertungen bei Probanden mit zunehmend umfangreicheren Nutzungserfahrungen.

Eine Detailbetrachtung der Ergebnisse des adaptiven Systems zeigte, dass negative Einschätzungen des adaptiven Testprodukts zu einem großen Teil mit gestalterischen Mängeln

der Umsetzung zusammenhängen. Diese Mängel verursachten Usability-Probleme und Einstiegshürden für die Nutzung. Die Notwendigkeit die Gestaltung von adaptiven Systemen zu verbessern, verdeutlichen speziell auch Ergebnisse der Abschlussbefragung nach Verbesserungspotentialen. Lösungsansätze, um diese Probleme zu minimieren, sind die Erstellung von Usability-Richtlinien für die Gestaltung adaptiver Systeme und die Klärung von Fragen zu adaptiver Funktionalität bereits bei der Anforderungsanalyse im Usability-Engineering Prozess.

Bei einer weiterführenden Studie sollten ergänzend auch Logg-Dateien oder spezifische Softwaretools eingesetzt werden um Usability-Probleme genauer bestimmen und die Gebrauchstauglichkeit eines adaptiven Produkts mittels objektiver quantitativer Daten untermauern zu können.

## 8 Quellen- und Literaturverzeichnis

Arnold, 1999

Arnold, A. G. (1999): Mental effort and evaluation of user-interfaces. A questionnaire approach. In: Bullinger, H.; Ziegler, J. (Hrsg.): Proceedings of the HCI International 1999, 1, S.1003-1007.

AttrakDiff, 2004

AttrakDiff Demoprojekt Ergebnisse. Untersuchungsbericht zum Produkt XY (2004): URL: [http://www.uidesign.de/attrakdiff/results/demo/demoprojekt\\_ergebnisse.pdf](http://www.uidesign.de/attrakdiff/results/demo/demoprojekt_ergebnisse.pdf), Stand: 18.10.2004.

Blom, 2000

Blom, J. (2000): Personalization. A taxonomy. Extended Abstracts of the CHI '00, S.313-314. New York: ACM.

Blom, 2002

Blom, J. (2002): A theory of personalized recommendations. Extended Abstracts of the CHI '02. New York: ACM.

Bolger, Davis, Rafaeli, 2002

Bolger, N.; Davis, A.; Rafaeli, E. (2002): Diary methods. Capturing life as it is lived. In: Annual Review of Psychology, 54, S.579-616.

Bortz, 1989

Bortz, J. (1989): Statistik für Sozialwissenschaftler. 3., neu bearbeitete Auflage. Berlin [u.a.]: Springer.

Bortz, Döring, 2003

Bortz, J.; Döring, N. (2003): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 3., überarbeitete Auflage. Berlin [u.a.]: Springer.

Bortz, Lienert, 2003

Bortz, J.; Lienert, G. A. (2003): Kurzgefasste Statistik für die klinische Forschung. Leitfaden für die verteilungsfreie Analyse kleiner Stichproben. 2., aktualisierte und bearbeitete Auflage. Berlin [u.a.]: Springer.

Burmester, 2003

Burmester, M. (2003): Unveröffentlichtes Dokument, München: Siemens AG.

Burmester, Wissmann, 2004

Burmester, M.; Wissmann, K. (2004): Unveröffentlichtes Dokument, München: Siemens AG.

Colbert, 2001

Colbert, M. (2001): A diary study of rendezvousing. Implication for position-aware computing and communications for the general public. In: Proceedings of the ACM 2001, S.15-23. New York: ACM.

DIN EN ISO-Norm 13407, 1999

DIN EN ISO 13407 (1999): Benutzerorientierte Gestaltung interaktiver Systeme.

DIN EN ISO-Norm 9241-10, 1996

DIN EN ISO 9241-10 (1996): Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten - Teil 10: Grundsätze der Dialoggestaltung.

DIN EN ISO-Norm 9241-11, 1998

DIN EN ISO 9241-11 (1998): Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten - Teil 11: Anforderungen an die Gebrauchstauglichkeit - Leitsätze.

Dubiella, 2003

Dubiella, S. (2003): Usability. Auswirkungen von personalisierten Dialog- und Menüstrukturen am Beispiel eines mobilen Endgeräts. Diplom-Arbeit. Universität Konstanz.

Eilers, Nachreiner, Hänecke, 1986

Eilers, K.; Nachreiner, F.; Hänecke, K. (1986): Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. In: Zeitschrift für Arbeitswissenschaft, 40, S.215-224.

Fast, 2002

Fast, K. (2002): Recording screen activity during usability testing. URL: [http://www.boxesandarrows.com/archives/recording\\_screen\\_activity\\_during\\_usability\\_testing.php](http://www.boxesandarrows.com/archives/recording_screen_activity_during_usability_testing.php), Stand: 18.10.2004.

Findlater, McGrenere, 2004

Findlater, L.; McGrenere, J. (2004): A comparison of static, adaptive, and adaptable menus. In: Papers of the CHI '04, 6, S.89-96.

Gibbs, 1997

Gibbs, A.(1997): Focus Groups. In: Social Research Update, 19, S.1-7.

Hamborg, Willumeit, Gediga, 1996

Hamborg, K.-C.; Willumeit, H.; Gediga, G. (1996): Untersuchungen zu Itemformulierungen des IsoMetrics-Verfahrens. Osnabrücker Schriftenreihe Software-Ergonomie, 1.

Hassenzahl, 2004

Hassenzahl, M. (2004): Unveröffentlichtes Dokument, München: Siemens AG.

Hassenzahl, Burmester, Koller, 2003

Hassenzahl, M.; Burmester, M.; Koller, F. (2003): AttrakDiff. Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Ziegler, J.; Szwillus, G. (Hrsg.): Mensch & Computer 2003. Interaktion in Bewegung, S.187-196. Stuttgart: Teubner.

Höök, 1997

Höök, K. (1997): Evaluating the utility and usability of an adaptive hypermedia system. In: IUI '97, S.179-186.

Horvitz, 1999

Horvitz, E. (1999): Principles of mixed-initiative user interfaces. In: Papers of the CHI '99, S.159-166.

Jameson, 2003

Jameson, A. (2003): Adaptive interfaces and agents. In: Jacko, J.; Sears, A. (Hrsg.): Human-computer action handbook, S.305-330. Mahwah, N.J.: Erlbaum.

Kantner, Sova, Rosenbaum, 2003

Kantner, L.; Sova, D. H.; Rosenbaum, S.(2003): Alternative methods for field usability research. In: Proceedings of the 21<sup>st</sup> annual international conference on documentation, S.68-72. New York: ACM.

Kobsa, 1993

Kobsa, A. (1993): Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen. In: Herzog, O.; Christaller, T.; Schütt, D. (Hrsg.): 17. Fachtagung KI. Berlin: Springer.

Kramer, Noronha, Vergo 2000

Kramer, J.; Noronha, S.; Vergo, J. (2000): A user-centered design approach to personalization. In: Communications of the ACM, 43, 8, S.45-48. New York: ACM.

Langley, 1997

Langley, P. (1997): Machine learning for adaptive user interfaces. In: Grewla, G.; Habel, C.; Nebel, B. (Hrsg.): KI '97. Advances in artificial intelligence, S.53-62.

Lanier, 1995

Lanier, J. (1995): Agents of alienation. In: Journal of Consciousness Studies, 2, 1, S.76-81.

Lewis, 1995

Lewis, J. R. (1995): IBM Computer Usability Satisfaction Questionnaires. Psychometric evaluation and instruction for use. In: International Journal of Human-Computer Interaction, 7, 1, S.57-78.

Mackay, 1991

Mackay, W. E. (1991): Triggers and barriers to customizing software. In: Proceedings of the SIGCHI conference on Human factors in computing systems '91, S.153-160. New York: ACM.

McGrenere, Baecker, Booth, 2002

McGrenere, J.; Baecker, R. M.; Booth, K. S. (2002): An evaluation of a multiple interface design solution for bloated software. In: Letters of the CHI '02, 4, 1, S.163-170.

McGrenere, Moore, 2000

McGrenere, J.; Moore, G. (2000): Are we all in the same „bloath“? In: Graphics Interface 2000, S.187-196.

Millen, 2000

Millen, D. R. (2000): Rapid ethnography. Time deepening strategies for HCI field research. In: Proceedings of the DIS '00, S.280-286. New York: ACM.

Nitschke, 2001

Nitschke, J. (2001): MTI-Workshop Nutzer- und Kontextadaptivität. URL: [http://www.dlr.de/PT-DLR/sw/workshop04/1\\_nitschke.pdf](http://www.dlr.de/PT-DLR/sw/workshop04/1_nitschke.pdf), Stand: 02.11.2004.

Norman, 1994

Norman, D. A. (1994): How might people interact with agents. In: Communications of the ACM, 37, 7, S.68-71. New York: ACM.

Parasuraman, Miller, 2004

Parasuraman, R.; Miller, C. A. (2004): Trust and ettiquette in high-criticality automated systems. In: Communications of the ACM, 47, 4, S.51-55. New York: ACM.

Preece, 1993

Preece, J. (1993): A guide to usability. Human factors in computing. Workingham: Addison-Wessley.

Prümper, 1997

Prümper, J. (1997): Der Benutzungsfragebogen ISONORM 9241/10. Ergebnisse zur Reliabilität und Validität. In: Liskowsky, R.; Velichkovsky, B. M.; Wünschmann, W. (Hrsg.): Software Ergonomie '97 - Usability Engineering. Integration von Mensch-Computer-Interaktion und Software-Entwicklung, S.253-262. Stuttgart: Teubner.

Riecken, 2000

Riecken, D. (2000): Personalized views of personalization. In: Communications of the ACM, 43, 8.

Rondoni, 2003

Rondoni, J. C. (2003): Context-aware experience sampling for the design and study of ubiquitous technologies. URL: [http://www.architecture.mit.edu/house\\_n/web/publications/thesis-documents/Rondoni03.pdf](http://www.architecture.mit.edu/house_n/web/publications/thesis-documents/Rondoni03.pdf), Stand 18.10.2004.

Shneiderman, 1995

Shneiderman, B. (1995): Looking for the bright side of user interface agents. In: Interactions, 2, 1, S.13-15. New York: ACM.

Weiss, 2002

Weiss, S. (2002): Handheld usability. Chichester: Wiley.

Wexelblatt, Maes, 2004

Wexelblatt, A.; Maes, P. (2004): Issues for software agent UI. URL: <http://web.media.mit.edu/~wex/agent-ui-paper/agent-ui.htm>, Stand: 18.10.2004.

Wissmann, 2003

Wissmann, K. (2003): Unveröffentlichtes Dokument, München: Siemens AG.

Zimmermann, 2004

Zimmermann, B. (2004): Ein Online-Internet-Kurs mit Glossar. Zum Einsteigen aber nicht für Einsteiger. URL: <http://www.www-kurs.de>, Stand: 18.10.2004.



## Anhänge

### Anhang A Rekrutierung

**S**

# Gesucht: Testpersonen!

**Studie Ergonomie von Office Software**

**Testobjekt:** Microsoft Word XP  
**Studienzeitraum:** Mitte Mai bis Ende Juli 2004 ( KW 22-30 = 24.05.-23.07.2004 )

**Benefiz:**

- Geschenk für Sie und Sekt-Knabberei / Kaffee-Kuchen als Dankeschön für Ihre gesamte Abteilung
- Verbesserungen im eigenen Umgang mit Office-Software
- Unterstützung der Siemens Forschungsarbeit



Detaillierte Informationen auf der Rückseite →

**S**

**Voraussetzung:** Sie arbeiten an Ihrem Arbeitsplatz seit kurzem mit einer deutschen Version von Microsoft Word XP bzw. in Ihrer Abteilung wird demnächst auf Microsoft XP umgestellt

**Sie müssen keine speziellen Tests durchführen!**

**Wir untersuchen Ihre Arbeitsgewohnheiten im Umgang mit Microsoft Word XP anhand Ihrer normalen Arbeitsaufgaben.**

Sie führen für uns ein **Tagebuch**. Der Aufwand liegt bei unter 15 min pro Woche!

An Ihrem Arbeitsplatz werden **3 Vor-Ort-Beobachtungen** à 2 h und je 1 Einstiegs- / Abschlusstreffen à 30 min stattfinden. Während der Vor-Ort-Beobachtungen werden alle Bewegungen auf Ihrem PC-Bildschirm mittels einer speziellen Software protokolliert.

- Wir geben keine Daten an Dritte weiter!
- Es wird für Sie keine Einschränkung während Ihrer Arbeit geben!
- Die erhobenen Daten dienen ausschließlich wissenschaftlichen Zwecken!



**Interesse??**

**Kontakt:**

<b>Kay Hauser</b> Diplomand (MSc)	<b>Siemens AG CT IC 7</b> Fon: +49 89 636- 54062 ctic7.project11@mchp.siemens.de
--------------------------------------	--

Abbildung 23 - Rekrutierung (Dateianhang, Aushang, Flugzettel)

## **Anhang B      Fragebogen zum Studienbeginn**

Studie „Ergonomie von Office-Software“

---

### **Herzlich Willkommen**

---

Willkommen zu der Studie „Ergonomie von Office-Software“.

Mit Ihrer Hilfe möchten wir herausfinden welche ergonomischen Mittel wir für eine benutzerfreundliche Gestaltung von Office-Produkten empfehlen können.

Dazu möchten wir mehr über Ihre Arbeitsgewohnheiten mit Microsoft Word XP erfahren. Die Untersuchung wird über eine Gesamtdauer von 9 Wochen vom 24.05.-23.07.2004 (Kalenderwochen 22 bis 30) stattfinden. Sollten Sie während dieses Zeitraums Urlaub genommen haben oder eine Zeit lang nicht in München sein, teilen Sie uns dies bitte mit.

Um eine größtmögliche Vergleichbarkeit der Ergebnisse zu gewährleisten, ist Voraussetzung um an der Untersuchung teilnehmen zu können, dass Sie mit der deutschen Version von Microsoft Word XP arbeiten.

Die gesamte Untersuchung wird auch in deutscher Sprache stattfinden.

Um uns vorab ein Bild von Ihnen machen zu können möchten wir Sie bitten uns noch ein paar Fragen zu beantworten.

Vielen Dank für Ihr Interesse.

## **Studienablauf**

---

**Sehr geehrte/r Studienteilnehmer/-in,**

das Fachzentrum User Interface Design der Siemens AG hat Sie um Teilnahme an der Studie „Ergonomie von Office-Software“ gebeten.

Ziel unserer Untersuchung ist die Beobachtung neuer Funktionalitäten in der XP Version von Microsoft Word. Es ist nicht das Ziel unserer Untersuchung, Sie als Testperson zu beobachten!

Für die Dauer der Studie möchten wir Sie bitten für uns ein Tagebuch zu führen. Der Aufwand liegt unter 15 Minuten pro Woche! Vorab möchten wir Sie informieren, dass wir vor Ort bei Ihnen am Arbeitsplatz Ihre Nutzungsgewohnheiten beobachten und Ihnen zu spezifische Ereignissen und Erfahrungen Fragen stellen. Während 3 Vor-Ort-Beobachtungen à 2 Stunden werden wir für etwa 90 Minuten eine spezielle Software einsetzen, die Aktivitäten auf Ihrem PC-Bildschirm protokolliert. Es wird für Sie hierbei nicht zu Einschränkungen Ihrer gewohnten Arbeit kommen! Außerdem werden wir ein Einstiegsgespräch und eine Abschlussbefragung von jeweils etwa 30 Minuten durchführen. Alle Daten sowie von Ihnen gemachte Aussagen in Fragebögen und Interviews werden wir selbst aus und verwenden sie zu rein wissenschaftlichen Zwecken! Wir geben keine Daten an Dritte weiter!

Wir werden an Ihrem Arbeitsplatz-PC an Microsoft Word XP spezifische Einstellungen vornehmen, welche Ihren normalen Arbeitsablauf jedoch nicht beeinträchtigen! Wir bitten Sie darum diese Einstellungen nicht zu verändern!

Somit helfen Sie uns und der Siemens AG benutzergerechte Produkte zu entwickeln.

## Fragebogen

---

Bitte ausfüllen... (Wenn Sie den Fragebogen elektronisch ausfüllen und uns per E-Mail senden möchten: Entsch die Kästchen anklicken und den Cursor in die grauen Textfelder setzen und Ihre Angaben eintippen)

### **Persönliche Angaben**

Name, Vorname :

Altersgruppe ☐ bis 29 ☐ 30 bis 39 ☐ 40 bis 54 ☐ ab 55

Wohnort :

Adressangabe (PC Standort)\* :

Telefonnummer :

E-Mail Adresse :

\* Bitte tragen Sie hier den Standort Ihres PCs ein, den Sie für die Studie verwenden möchten. Für Siemens Mitarbeiter z. B. den Siemens Standortplus Gebäude- und Raumnummer. Für private Studienteilnehmer Ihre normale Adresse.

### **Höchster Ausbildungsabschluss**

☐ Ausbildungsberuf, Volontariat :

☐ Weiterführende Ausbildung (Meister, Techniker, ...) :

☐ Hochschulstudium :

☐ Weiterführende akademische Ausbildung  
(Promotion, PhD-Programm, Habilitation, ...) :

☐ Andere / Sonstige :

Derzeitige Tätigkeit :

**PC-Kenntnisse**

Mehrfachauswahl möglich...

- ☐ Betriebssysteme :
- ☐ Datenbanken :
- ☐ Internet :
- ☐ Netzwerke :
- ☐ Office-Anwendungen :
- ☐ Programmierung :
- ☐ Sonstige Anwendungen :

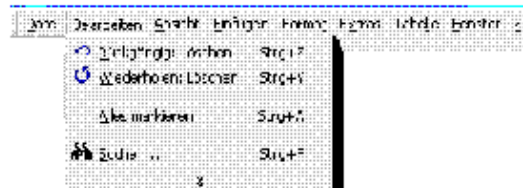
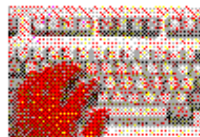
**Welche Microsoft Office Produktversion(en) verwenden Sie derzeit?**

Mehrfachauswahl möglich...

	Geschäftlich	Privat
Office 2003	<input type="checkbox"/>	<input type="checkbox"/>
Office XP	<input type="checkbox"/>	<input type="checkbox"/>
Office 2000	<input type="checkbox"/>	<input type="checkbox"/>
Office 97	<input type="checkbox"/>	<input type="checkbox"/>
Ältere Office Version	<input type="checkbox"/>	<input type="checkbox"/>

Wenn Sie mit Microsoft Office Produkten arbeiten welche Interaktionselemente verwenden Sie?

Mehrfachauswahl möglich...

☐ Menūs☐ Symbole☐ Tastatur-Shortcuts (z.B. Strg+S für Speichern)☐ Assistant

***Würden Sie sich wünschen, dass Microsoft Office Produkte Ihren individuellen Arbeitsgewohnheiten angepasst sind?***

---

Mehrfachauswahl möglich...

- ☐ Ja, sollten sich automatisch anpassen, ...
- ☐ ...auch ohne mein Zutun.
- ☐ ...aber nur wenn ich informiert werde über die vorgenommenen Anpassungen.
- ☐ ...aber nur mit meiner ausdrücklichen Zustimmung.
- ☐ Ja, durch die Möglichkeit, selbst Anpassungen vornehmen zu können.
- ☐ Nein, ich bin zufrieden mit der Standardversion.

***Wie häufig nutzen Sie Microsoft Word?***

---

- ☐ Täglich, über 4 Stunden
- ☐ Täglich, aber unter 4 Stunden
- ☐ Mehrmals wöchentlich
- ☐ Nicht mehr als einmal pro Woche

***Wenn Sie mit Microsoft Word arbeiten welche Art von Aufgaben führen Sie aus?***

---

Beispiele für Aufgaben wären etwa das Schreiben eines Briefs, die Anfertigung einer Übersetzung oder die graphische Gestaltung eines Handouts,

Meistgleiche Aufgaben      ☐    ☐    ☐    ☐    ☐    Meist wechselnde Aufgaben

## Anhang C    *Tagebuch*

### Kurze Einführung in das Tagebuch

---

Einen Eintrag in das Tagebuch nehmen Sie bitte dann vor, wenn Sie eine Tätigkeit mit Microsoft Word XP abgeschlossen, also eine **Aufgabe erledigt** haben. Aufgaben wären z.B. das Schreiben eines Briefes, die Erstellung eines Angebots oder das Anfertigen einer Übersetzung, aber auch das Lesen eines Dokuments.

Wenn Sie eine Tätigkeit ausführen, die einen längeren Zeitraum in Anspruch nimmt, wie etwa das Schreiben eines umfangreicheren Aufsatzes, würde diese **Tätigkeit nach jeder abgelaufenen Stunde als eigenständige Aufgabe** gewertet werden.

**Pro Tag** brauchen Sie jedoch **nicht mehr als zwei bis drei Aufgaben** im Tagebuch zu notieren. Am besten Sie machen immer für die ersten Aufgaben eines Tages einen Eintrag. Sollten Sie einmal vergessen einen Eintrag für eine Aufgabe zu machen, brauchen Sie diesen nicht nachträglich vorzunehmen, machen Sie stattdessen einfach einen Eintrag für die nächste Aufgabe.

Das Tagebuch ist aufgebaut wie ein kurzer Fragebogen. Die **Fragen beziehen sich auf die Aufgabe**, die Sie ausgeführt haben, und auf Ihr **Arbeiten mit den Menüs** bei Microsoft Word XP. Besonders interessant sind für uns dabei **Ereignisse**, die Sie bei Ihrer Arbeit unterstützen oder behindern.

Füllen Sie das **Tagebuch** bitte immer **vollständig** aus.

Sollten Sie einmal Probleme oder Kommentare zum Tagebuch haben, teilen Sie uns dies in der Abschlussfrage „Hatten Sie Probleme die Fragen zu beantworten? Oder haben Sie weitere Kommentare zu dem Tagebuch“ bitte mit. So helfen Sie uns die Vorgehensweise zu optimieren.

---



### ***Aufgabenbeschreibung***

---

**Bitte beschreiben Sie kurz die Aufgabe, die Sie ausgeführt haben:**

Zum Beispiel: das Schreiben eines Briefes oder das Anfertigen einer Übersetzung.

**Dauer der Aufgabe:**

- ☐ Etwa 1 Stunde oder länger
- ☐ Über 15 Minuten, aber unter 1 Stunde
- ☐ Unter 15 Minuten

**Bitte beurteilen Sie die Schwierigkeit Ihrer Aufgabe:**

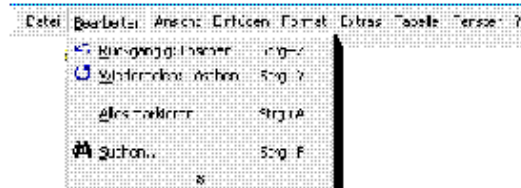
Mit der Schweregrad Ihrer Aufgabe ist hier gemeint, wie einfach oder kompliziert Sie z.B. das Anfertigen einer englischen Übersetzung einschätzen aufgrund Ihrer englischen Sprachkenntnisse. Nicht gemeint ist hingegen, wie schwierig Ihnen die Erledigung dieser Aufgabe mit Microsoft Word XP gefallen ist!

Einfach ☐ ☐ ☐ ☐ ☐ ☐ ☐ Kompliziert

**Welche Interaktionselemente haben Sie bei der Bearbeitung Ihrer Aufgabe verwendet?**

Mehrtauswahl möglich...

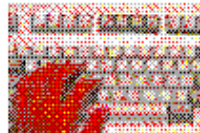
☐ Menüs



☐ Symbole



☐ Tastatur-Shortcuts (z.B. Strg+S für Speichern)



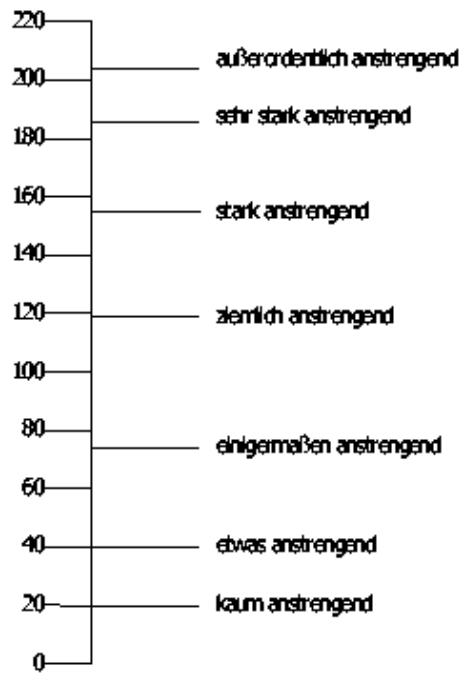
☐ Assistent



### **Anstrengung**

---

Geben Sie bitte auf der untenstehenden senkrechten Linie an - mit einem Kreuzchen oder Strich - wie anstrengend Sie das Ausführen der Aufgabe mit den Menüs von Microsoft Word XP empfunden haben.



### ***Welche Ereignisse traten während der Bearbeitung der Aufgabe mit den Menüs auf?***

Beim Arbeiten mit den Menüs von Microsoft Word XP treten Ereignisse auf, die Sie bei der Erledigung Ihrer Aufgabe unterstützen oder behindern. Bitte geben Sie an welche Ereignisse bei der Erledigung Ihrer Aufgabe aufgetreten sind. Die Ereignisse sehen wie folgt aus:

#### **Unterstützende Ereignisse**

- ☐ Ich konnte den **Menü**eintrag schnell dort finden, wo ich ihn **erwartet** hatte.
- ☐ Ich konnte den Menüeintrag schnell finden, da das **Menü** kurz und **übersichtlich** war.
- ☐ Ich konnte den **Menü**eintrag schnell finden, da er **prominent** platziert war.

Wie oft sind diese Ereignisse bei der Bearbeitung Ihrer Aufgabe aufgetreten?	Nie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Immer
--	-----	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------

Haben diese Ereignisse Sie bei der Bearbeitung Ihrer Aufgabe unterstützt?	Gar nicht	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Außerordentlich
---	-----------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-----------------

#### **Behindernde Ereignisse**

- ☐ Ich konnte den **Menü**eintrag nicht dort finden, wo ich ihn **erwartet** hatte.
- ☐ Ich musste den Menüeintrag erst suchen, da das **Menü** lang und **unübersichtlich** war.
- ☐ Ich konnte den **Menü**eintrag gar nicht finden. Stattdessen bin ich auf ein anderes Interaktionselement (Symbole, Shortcuts, Assistent) ausgewichen.
- ☐ Um den Menüeintrag zu finden, musste ich das **Menü** verlängern.

Wie oft sind diese Ereignisse bei der Bearbeitung Ihrer Aufgabe aufgetreten?	Nie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Immer
--	-----	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------

Haben diese Ereignisse Sie bei der Bearbeitung Ihrer Aufgabe behindert?	Gar nicht	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Außerordentlich
---	-----------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-----------------

### ***Bewertung der aufgetretenen Ereignisse bei den Menüs von Microsoft Word XP***

---

Nachfolgend finden Sie Wortpaare, mit deren Hilfe Sie die, in der vorigen Frage geschilderten, bei den Menüs von Microsoft Word XP aufgetretenen Ereignisse bewerten können. Die Wortpaare stellen jeweils extreme Gegensätze dar, zwischen denen eine Abstufung möglich ist.

Die Darstellung / Veranschaulichung der Ereignisse empfand ich als...

Angenehm	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Störend
----------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------

Aufgrund der Ereignisse sehe ich meine Privatsphäre durch das System...

Gewahrt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Gefährdet
---------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-----------

Generell ist das Auftreten von Ereignissen...

Vorhersagbar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Unberechenbar
--------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------------

Das Systemverhalten empfinde ich als...

Zuverlässig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Unzuverlässig
-------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------------

Meine Bedürfnisse werden vom System...

Verstanden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nicht verstanden
------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	------------------

Ich fühle mich dem System gegenüber...

In Kontrolle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Ausgeliefert
--------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------

### ***Hatten Sie Probleme die Fragen zu beantworten? Oder haben Sie weitere Kommentare zu dem Tagebuch?***

---

Wenn Sie an einer Stelle des Tagebuchs Probleme hatten oder vielleicht etwas vermisst haben oder einfach einen Kommentar zu dem Tagebuch abgeben möchten, können Sie uns dies hier gerne mitteilen.

So helfen Sie uns die Vorgehensweise bei künftigen Studien noch zu optimieren. Vielen Dank.

## Anhang D    *Protokoll zur Vor-Ort-Beobachtung*

Vor-Ort-Beobachtung – Intervention-Protokoll  
[ ausgefüllt vom Studienleiter ]

**SIEMENS**

Name der Testperson:

Vor-Ort-Beobachtung, durchgeführt am

Aufgabe Nummer:

Startzeit:

**Dauer der Aufgabe:**

- ☐ Etwa 1 Stunde oder länger  
☐ Über 15 Minuten, aber unter 1 Stunde  
☐ Unter 15 Minuten

**Verwendete Interaktionselemente:**

- ☐ Menüs  
☐ Symbole  
☐ Tastatur-Shortcuts  
☐ Assistent

**Beobachtete Ereignisse Übersicht:**

- |                          |  |
|--------------------------|--|
| <input type="checkbox"/> | Eintrag wo erwartet                    |
| <input type="checkbox"/> | Dropdown-Menü kurz und übersichtlich   |
| <input type="checkbox"/> | Eintrag prominent platziert            |
| <input type="checkbox"/> | Eintrag nicht wo erwartet              |
| <input type="checkbox"/> | Dropdown-Menü lang und unübersichtlich |
| <input type="checkbox"/> | Menüeintrag gar nicht gefunden         |
| <input type="checkbox"/> | Dropdown-Menü verlängert               |

**Andere Beobachtungen:**

Uhrzeit	Beobachtetes Ereignis	Beobachtete Reaktion Proband
	<input type="checkbox"/> Eintrag wo erwartet <input type="checkbox"/> Dropdown-Menü kurz und übersichtlich <input type="checkbox"/> Eintrag prominent platziert  <input type="checkbox"/> Eintrag nicht wo erwartet <input type="checkbox"/> Dropdown-Menü lang und unübersichtlich <input type="checkbox"/> Menüeintrag gar nicht gefunden <input type="checkbox"/> Dropdown-Menü verlängert	
	<input type="checkbox"/> Eintrag wo erwartet <input type="checkbox"/> Dropdown-Menü kurz und übersichtlich <input type="checkbox"/> Eintrag prominent platziert  <input type="checkbox"/> Eintrag nicht wo erwartet <input type="checkbox"/> Dropdown-Menü lang und unübersichtlich <input type="checkbox"/> Menüeintrag gar nicht gefunden <input type="checkbox"/> Dropdown-Menü verlängert	



## **Anhang E      Fragebogen zur Intervention**

Studie „Ergonomie von Office-Software“ – Fragebogen Vor-Ort-Beobachtung

---

### **Fragebogen-Instruktionen**

---

**Sehr geehrte/r Studienteilnehmer/-in,  
bitte unbedingt lesen!**

Im Folgenden geht es um Ihre individuelle Erfahrung während der Arbeit mit Microsoft Word XP seit unserem letzten Treffen.

Dabei geht es **nicht** um eine Beurteilung Ihrer Person, sondern um Ihre persönliche Einschätzungen zur Arbeit mit Microsoft Word XP.

**Lassen Sie bitte keine der Fragen aus!**

## Fragebogen

Bitte ausfüllen...

### Beurteilung von Microsoft Word XP

Nachfolgend finden Sie Wortpaare, mit deren Hilfe Sie Microsoft Word XP bewerten können. Die Wortpaare stellen jeweils extreme Gegensätze dar, zwischen denen eine Abstufung möglich ist.

Ein Beispiel:

gut ☐ ☐ ☐ ☐ ☐ ☒ ☐ schlecht

Diese Bewertung bedeutet, dass Sie Microsoft Word XP eher schlecht bewerten.

Denken Sie nicht lange über die Wortpaare nach, sondern geben Sie bitte die Einschätzung ab, die Ihnen spontan in den Sinn kommt.

Vielleicht passen einige Wortpaare nicht so gut auf das Produkt, kreuzen Sie aber trotzdem bitte immer eine Antwort an. Denken Sie daran, dass es keine "richtigen" oder "falschen" Antworten gibt - nur Ihre persönliche Meinung zählt!

menschlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	technisch
isolierend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	verbindend
angenehm	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unangenehm
originell	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	konventionell
einfach	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	kompliziert
fachmännisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	laienhaft
hässlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	schön
praktisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unpraktisch
sympathisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unsympathisch
umständlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	direkt
stilvoll	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	stillos
voraussagbar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unberechenbar
minderwertig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	wertvoll
ausgrenzend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	einbeziehend
bringt mich den Leuten näher	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	trennt mich von Leuten
nicht vorzeigbar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	vorzeigbar
zurückweisend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	einladend
phantasielos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	kreativ
gut	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	schlecht

verwirrend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	übersichtlich
abstoßend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	anziehend
mutig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	vorsichtig
innovativ	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	konservativ
lahm	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	fesselnd
harmlos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	herausfordernd
motivierend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	entmutigend
neuartig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	herkömmlich
widerspenstig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	handhabbar

---

### Aufgabenbearbeitung

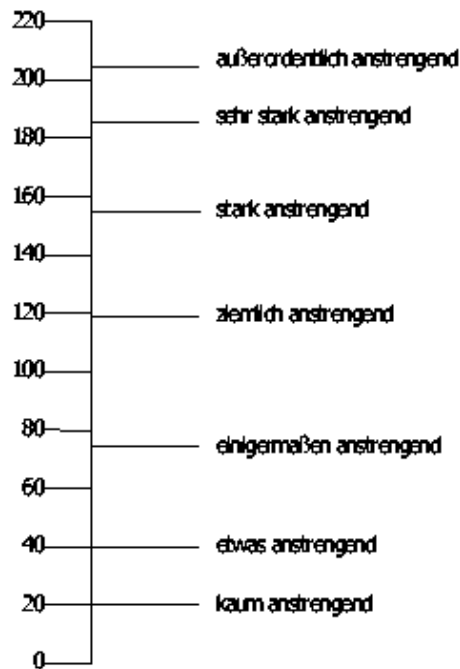
Wie würden Sie im Mittel die Schwierigkeit der Aufgabe einschätzen, die Sie seit unserem letzten Treffen ausgeführt haben:

Gemeint ist hier, wie schwer es Ihnen fällt z.B. eine englische Übersetzung anzuliegen aufgrund Ihrer englischen Sprachkenntnisse. ~~Wird~~ gemeint ist hingegen, wie schwer Ihnen die Erledigung der Aufgabe mittels Microsoft Word XP gefallen ist!

Einfach ☐ ☐ ☐ ☐ ☐ ☐ ☐ Kompliziert

### Anstrengung

Geben Sie bitte auf der unterstehenden senkrechten Linie an - mit einem Kreuzchen oder Strich - wie anstrengend Sie das Ausführen der Aufgaben mit den ~~Werten~~ von Microsoft Word XP empfunden haben.



### Arbeiten mit den Menüs von Microsoft Word XP

Hier möchten wir Sie bitten sich Gedanken zu dem Arbeiten mit den **Menüs** von Microsoft Word XP zu machen und diese zu bewerten.

#### Effizienz

---

Mit den <b>Menüs</b> von Microsoft Word XP kann ich meine Aufgaben effizient erledigen	Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig
--	---------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	----------------

#### Erlernbarkeit

---

Im Folgenden möchten wir wissen, wie gut das Arbeiten mit den **Menüs** von Microsoft Word XP für Sie erlernbar ist.

Das Arbeiten mit den **Menüs** von Microsoft Word XP...

Erfordert viel Zeit zum Erlernen.	Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig
Ermutigt dazu, auch mal neue Funktionen auszuprobieren.	Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig
Erfordert, dass man sich viele Details merken muss.	Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig
Ist so gestaltet, dass sich einmal Gelerntes gut einprägt	Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig

### Bewertung der aufgetretenen Ereignisse bei den Menüs von Microsoft Word XP

Bei dieser Frage geht es um die aufgetretenen Ereignisse bei den Menüs von Microsoft Word XP, entsprechend den Einträgen in das Tagebuch. Diese Ereignisse sehen wie folgt aus:

#### Unterstützende Ereignisse

- Ich konnte den gesuchten Menüeintrag schnell dort finden, wo ich ihn erwartete.
- Ich konnte den Menüeintrag schnell finden, da das Menü klar und übersichtlich war.
- Ich konnte den Menüeintrag schnell finden, da er prominent platziert war.

#### Behindernde Ereignisse

- Ich konnte den Menüeintrag nicht dort finden, wo ich ihn erwartete.
- Ich musste den Menüeintrag erst suchen, da das Menü lang und unübersichtlich war.
- Ich konnte den Menüeintrag gar nicht finden. Stattdessen bin ich auf ein anderes Interaktionselement (Symbole, Shortcuts, Assistent) ausgewichen.
- Um den Menüeintrag zu finden, musste ich das Menü verlängern.

Nachfolgend finden Sie Wortpaare, mit deren Hilfe Sie die diese Ereignisse bewerten können. Die Wortpaare stellen jeweils extreme Gegensätze dar, zwischen denen eine Abstufung möglich ist.

Die Ereignisse haben mich bei der Erledigung meiner Aufgabe...

Unterstützt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Behindert
-------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-----------

Die Darstellung / Veranschaulichung der Ereignisse empfand ich als...

Angenehm	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Störend
----------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------

Aufgrund der Ereignisse sehe ich meine Privatsphäre durch das System...

Gewahrt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Gefährdet
---------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-----------

Generell ist das Auftreten von Ereignissen...

Vorhersagbar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Unberechenbar
--------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------------

Das Systemverhalten empfinde ich als...

Zuverlässig	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Unzuverlässig
-------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------------

Meine Bedürfnisse werden vom System...

Verstanden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nicht verstanden
------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	------------------

Ich fühle mich dem System gegenüber...

In Kontrolle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Ausgeliefert
--------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------

***Hatten Sie Probleme die Fragen zu beantworten? Oder haben Sie weitere Kommentare zu dem Fragebogen?***

---

Wenn Sie an einer Stelle des Fragebogens Probleme hatten oder vielleicht etwas vermisst haben oder einfach einen Kommentar zu dem Fragebogen abgeben möchten, können Sie uns das hier gerne mitteilen.

So helfen Sie uns die Vorgehensweise bei künftigen Studien noch zu optimieren. Vielen Dank.

**Vielen Dank!**

## **Anhang F     Recall Survey Protokoll**

### **Recall Survey Protokoll zur**

Name des Studienteilnehmers:                      Testgruppe:

**Studienphase**                      (vom                      bis                      )

**Ergebnisse der Tagebuchsichtung:**

Wichtige Bemerkungen zur Tagebuchführung:

Unvollständige Tagebucheinträge:

Aufgaben Gesamt:

Aufgaben mit Verwendung der Menüs:

Schwierigkeit / Anstrengung:

Unterstützenden Ereignisse:

Behindemde Ereignisse:

Bewertungen:

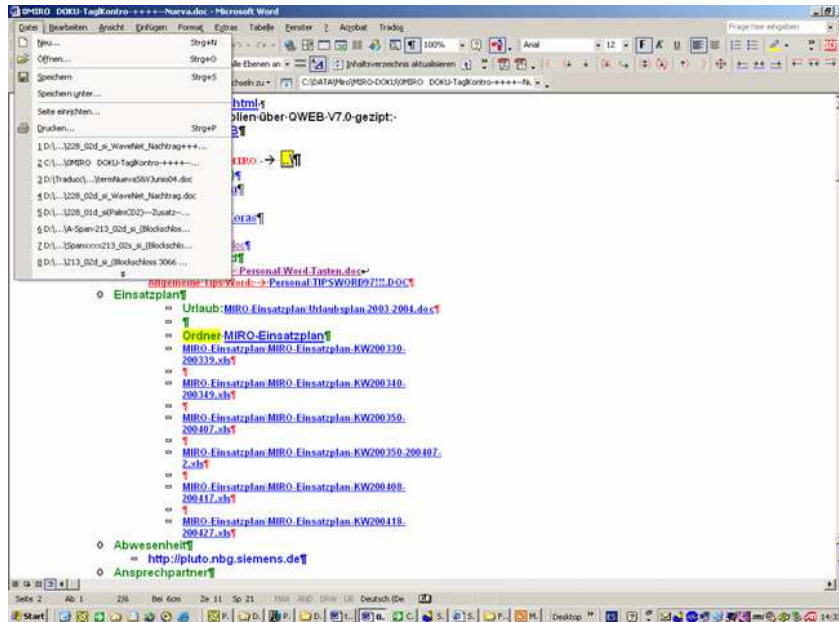
Auffälligkeiten:

Analyse-Ergebnisse:

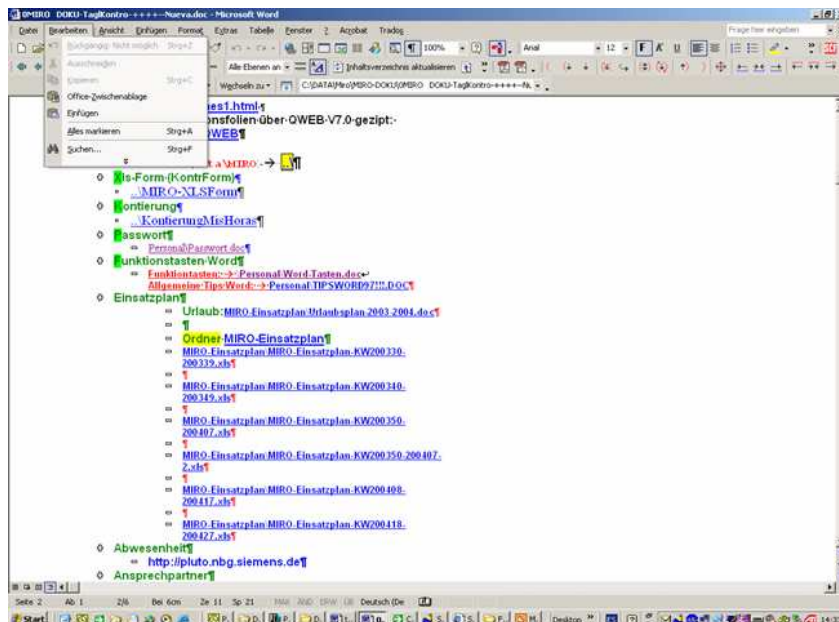
Kommentare der Testperson:



Auszug der Dokumentation der Benutzungsoberfläche vom 07.06.2004, Screenshots eines Probanden der Adaptiven Gruppe



**Abbildung 24 - Screenshot Dokumentation der Benutzeroberfläche Beispiel 1**



**Abbildung 25 - Screenshot Dokumentation der Benutzeroberfläche Beispiel 2**

## **Anhang H      Fragebogen zum Studienabschluss**

Studie „Ergonomie von Office-Software“ – Abschlussbefragung

---

### **Verbesserung von Microsoft Word XP**

Hier möchten wir Sie bitten sich Gedanken zur Verbesserung von Microsoft Word XP zu machen.

**Würden Sie sich wünschen, dass Microsoft Word XP Ihren individuellen Arbeitsgewohnheiten angepasst ist?**

---

Mehrfachauswahl möglich...

- ☐ Ja, sollte sich automatisch anpassen, ...
  - ☐ ...auch ohne mein Zutun.
  - ☐ ...aber nur wenn ich informiert werde über die vorgenommenen Anpassungen.
  - ☐ ...aber nur mit meiner ausdrücklichen Zustimmung.
- ☐ Ja, durch die Möglichkeit, selbst Anpassungen vornehmen zu können.
- ☐ Nein, ich bin zufrieden mit der Standardversion.

**Welche Verbesserungen würden Sie sich bei Microsoft Word XP generell wünschen?**

---

Bitte erläutern Sie kurz mit Ihren eigenen Worten...

**Welche Verbesserungen speziell im Bezug auf die Menüs bei Microsoft Word XP würden Sie sich wünschen?**

Nachfolgend finden Sie Aussagen zur Verbesserung der Menüs von Microsoft Word XP. Bitte schätzen Sie Ihre Zustimmung zu jeder Aussage auf der unter der Frage befindlichen Skala ein. Für den Fall, dass Sie sich aus irgendwelchen Gründen dazu nicht äußern wollen oder können, sollten Sie "keine Angabe" ankreuzen.

Die Handhabung der Menüs von Microsoft Word XP sollte einfacher sein.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Die Menüs von Microsoft Word XP sollten in sich stimmiger / widerspruchsfreier sein.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Die Menüs von Microsoft Word XP sollten leichter kontrollierbar sein.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Das Arbeiten mit den Menüs von Microsoft Word XP sollte mehr Spaß machen.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Die Menüs von Microsoft Word XP sollten mich bei der Erledigung meiner Aufgaben besser unterstützen.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Die Menüs von Microsoft Word XP sollten besser auf meine individuellen Wünsche und Aufgaben abstimbar / abgestimmt sein.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Automatische Anpassungen sowie die Möglichkeit, bei den Menüs selbst Anpassungen vornehmen zu können, sollten besser in Microsoft Word XP integriert sein.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Einstellungen zu automatischen Anpassungen oder der Möglichkeit, bei den Menüs selbst Anpassungen vornehmen zu können, sollten bei Microsoft Word XP einfacher zugänglich sein.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Das Verhalten der Menüs von Microsoft Word XP sollte zuverlässiger sein.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Ich wünsche mir verständlichere, besser visualisierte und angenehmer gestaltete Rückmeldungen zu automatischen Anpassungen der Menüs von Microsoft Word XP.

Völlig Falsch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Völlig richtig	Keine Angabe	<input type="checkbox"/>
------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------	-----------------	--------------------------

Wenn Sie noch weitere Aussagen zur Verbesserung der Menüs von Microsoft Word XP machen möchten, können Sie dies hier gerne tun.

Bitte erläutern Sie kurz mit Ihren eigenen Worten...

---

### Feedback zum Studienablauf

Wenn Sie an einer Stelle der Studie „Ergonomie von Office-Software“ Probleme hatten oder vielleicht etwas vermisst haben oder einfach einen Kommentar abgeben möchten, können Sie uns dies hier gerne mitteilen.

So helfen Sie uns die Vorgehensweise bei künftigen Studien noch zu optimieren. Vielen Dank.

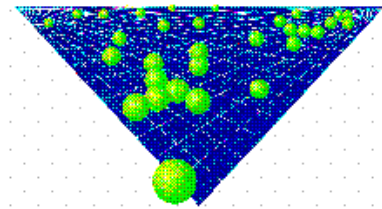
**Vielen Dank!**

*Anhang I      Protokoll des Interviews zum Studienabschluss*

**SIEMENS**

Protokoll zum Interview der Abschlussbefragung  
[ ausgefüllt vom Studienleiter ]

**Studie „Ergonomie von Office Software“**



Name der Testperson:

Testgruppe:

Abschlussbefragung, durchgeführt am

Angaben der Testperson zum Studienbeginn:

Ermittelte Einschätzung während der Studie:

### **Fragen zur Arbeitsweise**

Einschätzung aufgrund der Vor-Ort-Beobachtungen:

Wie beschreibt der Studienteilnehmer seine eigene Arbeitsweise:

Wann greift er auf welches Interaktionselement zurück?

Was gefällt / stört ihn an welchem Interaktionselement?

Was ist ihm während der Studie an bestimmten Interaktionselementen aufgefallen?

- ☐ Menüs
- ☐ Symbole
- ☐ Tastatur-Shortcuts
- ☐ Assistent
- ☐ Kontextuelle Interaktionselemente

### **Fragen zu Personalisierung und Adaptivität**

Systemeinstellung vor der Studie: ☐ Adaptiv ☐ Non ☐ XP Neu

#### **Wahrnehmen von Anpassungen bei Menüs**

[ NUR EXPERIMENTALGRUPPE!! ]

☐ Proband war der Kontrollgruppe zugeordnet

Hat der Studienteilnehmer Anpassungen während der Studie wahrgenommen?

Hat der Studienteilnehmer eventuell wahrgenommene Anpassungen mit den Fragen zu den definierten Ereignissen im Tagebuch, Fragebogen in Verbindung gebracht?

Ist er sich im Klaren, dass diese Anpassungen seine Arbeitsweise reflektieren und ihn unterstützen sollen?

☐ Ja

☐ Nein

Wie sieht er diese Anpassungen in seinen eigenen Worten?

#### **Personalisierung von Interaktionselementen bei Microsoft Word XP:**

Wunsch während der Studie geäußert, bereits personalisiert, Wunsch nach Ende der Studie geäußert...

☐ Nein

☐ Ja, welche...

☐ Menüs

☐ Symbole

☐ Tastatur-Shortcuts

Aus welchen Gründen hat die Testperson diese Anpassungen vorgenommen / würde sie vornehmen wollen?

Welche Anpassungen im Detail hat die Person vorgenommen / würde sie vornehmen wollen?

## **Bewertungen**

### **Zusammenfassung**

Tagebuch

Fragebogen

Auffälligkeiten

Fragen zur Bewertung an den Studienteilnehmer

Weitere Vorkommnisse



## **Anhang J**     ***Tücken der Datenerhebung: Die multitechnologische Umsetzung des BSMA im Tagebuch***

**Anstrengung**

Geben Sie bitte anhand der Kontrollkästchen links der Skala an, wie anstrengend Sie das Ausführen der Aufgabe mit den Menüs von Microsoft Word XP empfunden haben.

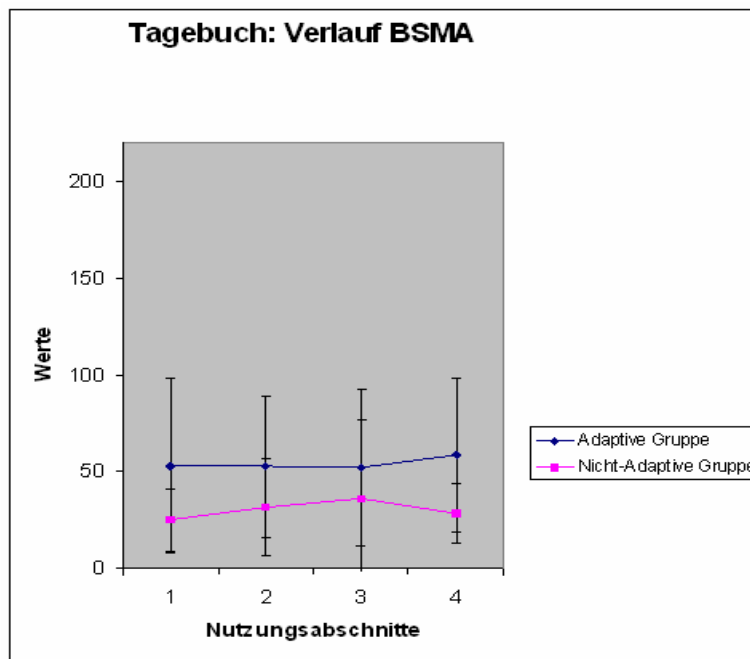
<input checked="" type="radio"/>	220	
<input type="radio"/>	200	außerordentlich anstrengend
<input type="radio"/>	180	sehr stark anstrengend
<input type="radio"/>	160	stark anstrengend
<input type="radio"/>	140	
<input type="radio"/>	120	ziemlich anstrengend
<input type="radio"/>	100	
<input type="radio"/>	80	einigermaßen anstrengend
<input type="radio"/>	60	
<input type="radio"/>	40	etwas anstrengend
<input type="radio"/>	20	kaum anstrengend
<input type="radio"/>	0	

### **Anstrengung**

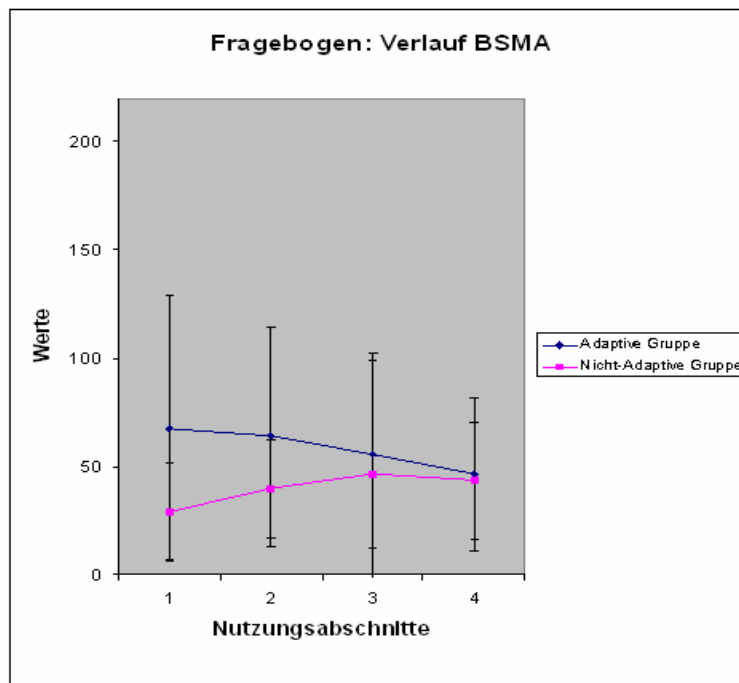
Geben Sie bitte auf der untenstehenden senkrechten Linie oder anhand der Kontrollkästchen links der Skala an, wie anstrengend Sie das Ausführen der Aufgabe mit den Maüs von Microsoft Word XP empfunden haben.

<input type="checkbox"/>	220	
<input type="checkbox"/>		
<input type="checkbox"/>	200	außerordentlich anstrengend
<input type="checkbox"/>		
<input type="checkbox"/>	180	sehr stark anstrengend
<input type="checkbox"/>		
<input type="checkbox"/>	160	stark anstrengend
<input type="checkbox"/>		
<input type="checkbox"/>	140	
<input type="checkbox"/>		
<input type="checkbox"/>	120	ziemlich anstrengend
<input type="checkbox"/>		
<input type="checkbox"/>	100	
<input type="checkbox"/>		
<input type="checkbox"/>	80	einigemaßen anstrengend
<input type="checkbox"/>		
<input type="checkbox"/>	60	
<input type="checkbox"/>		
<input type="checkbox"/>	40	etwas anstrengend
<input type="checkbox"/>		
<input type="checkbox"/>	20	kaum anstrengend
<input type="checkbox"/>		
<input type="checkbox"/>	0	

**Anhang K**     *Dritte Forschungsfrage: Ergebnisse der Erhebungsinstrumente BSMA und Anforderungen an adaptive Systeme*



**Abbildung 26 - Dritte Forschungsfrage: Ergebnis BSMA Tagebuch**



**Abbildung 27 - Dritte Forschungsfrage: Ergebnis BSMA Fragebogen**

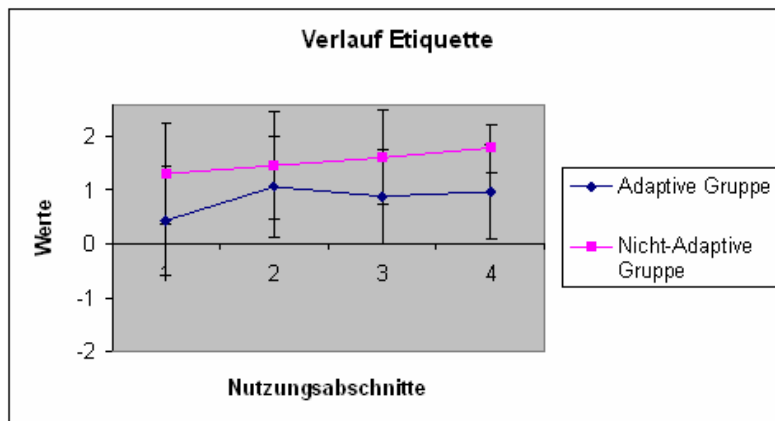


Abbildung 28 - Dritte Forschungsfrage: Ergebnis Etiquette Tagebuch

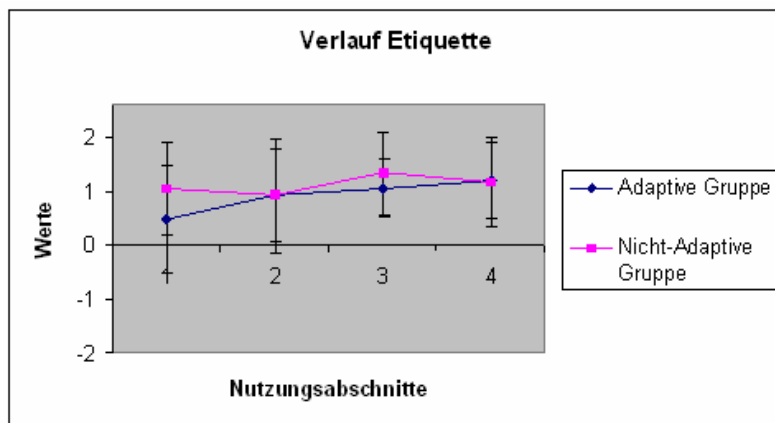


Abbildung 29 - Dritte Forschungsfrage: Ergebnis Etiquette Fragebogen

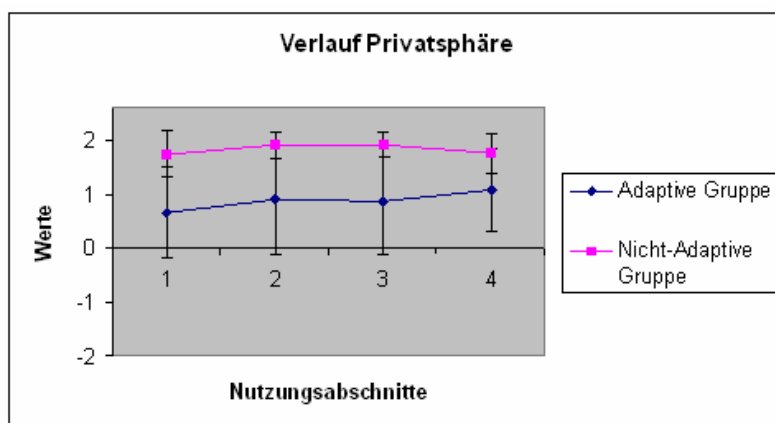


Abbildung 30 - Dritte Forschungsfrage: Ergebnis Privatsphäre Tagebuch

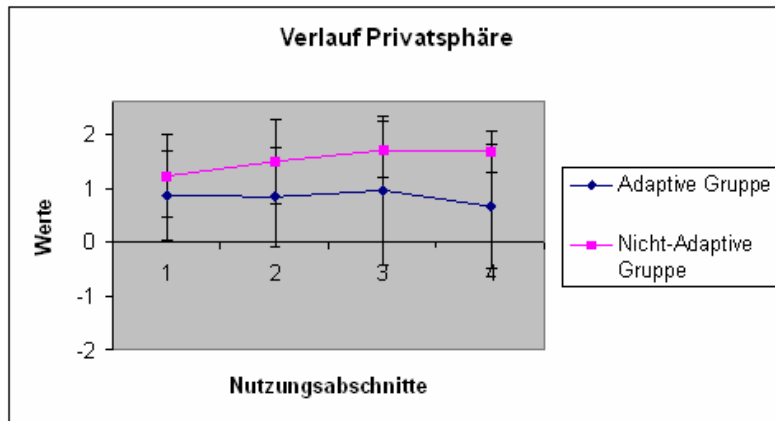


Abbildung 31 - Dritte Forschungsfrage: Ergebnis Privatsphäre Fragebogen

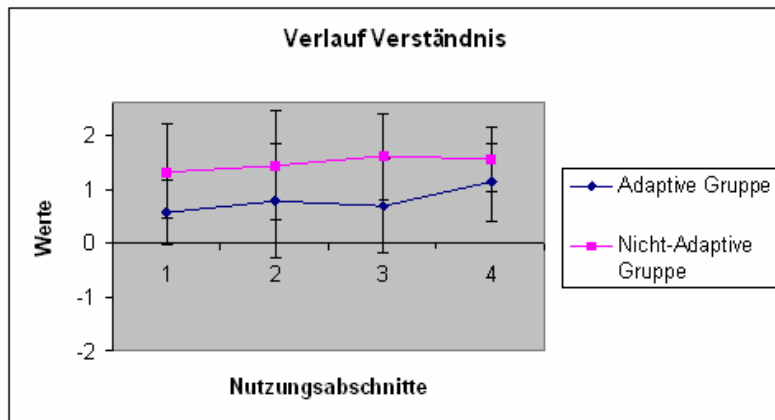


Abbildung 32 - Dritte Forschungsfrage: Ergebnis Verständnis Tagebuch

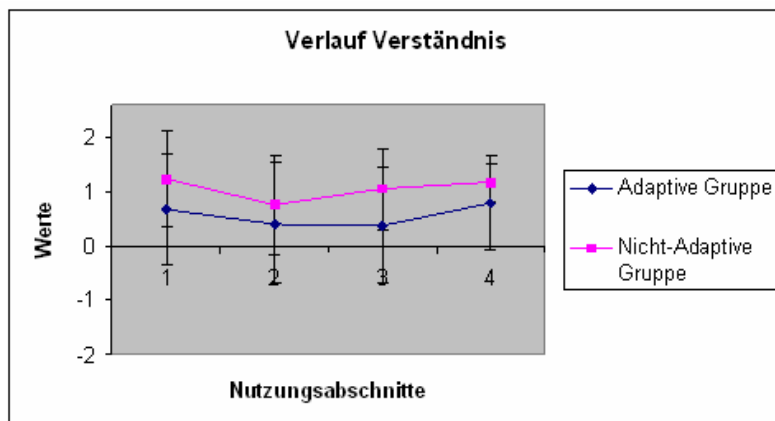


Abbildung 33 - Dritte Forschungsfrage: Ergebnis Verständnis Fragebogen

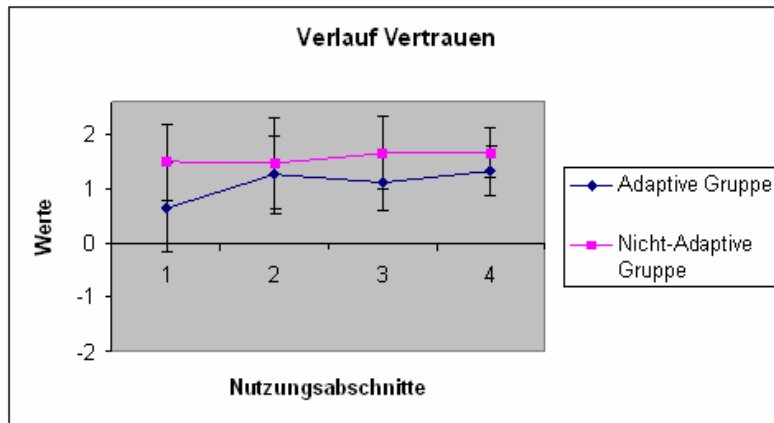


Abbildung 34 - Dritte Forschungsfrage: Ergebnis Vertrauen Tagebuch

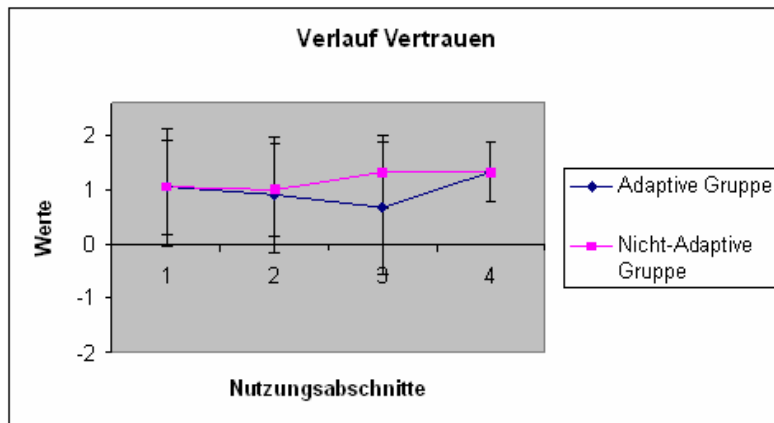


Abbildung 35 - Dritte Forschungsfrage: Ergebnis Vertrauen Fragebogen

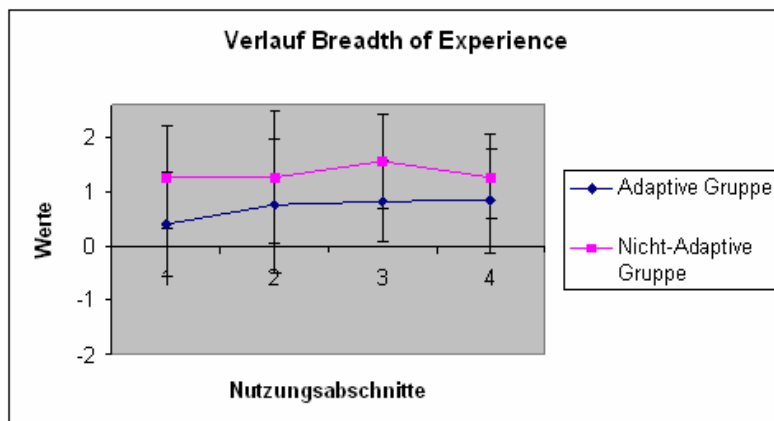


Abbildung 36 - Dritte Forschungsfrage: Ergebnis Breadth of Experience Tagebuch

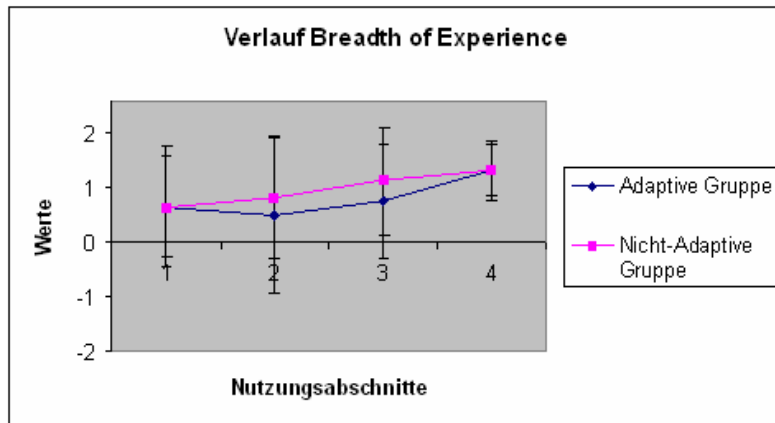


Abbildung 37 - Dritte Forschungsfrage: Ergebnis Breadth of Experience Fragebogen

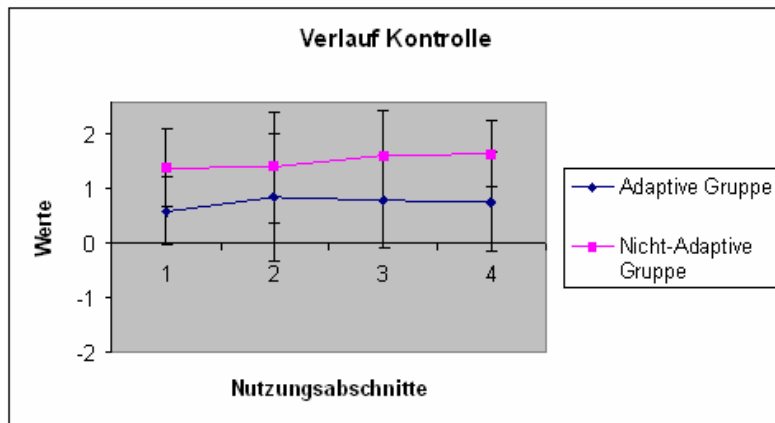


Abbildung 38 - Dritte Forschungsfrage: Ergebnis Kontrolle Tagebuch

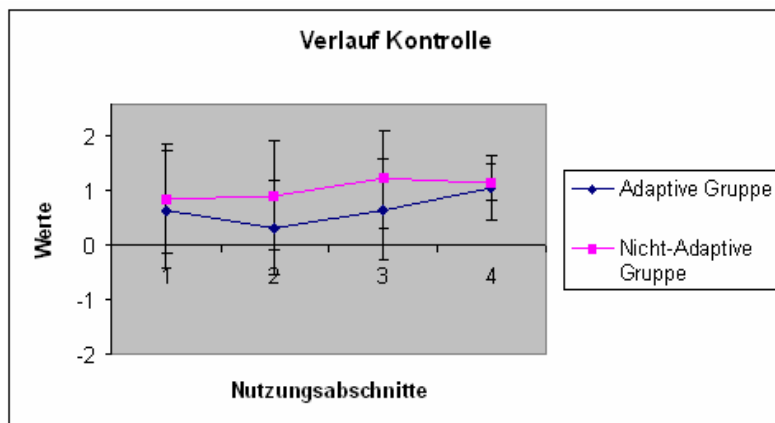
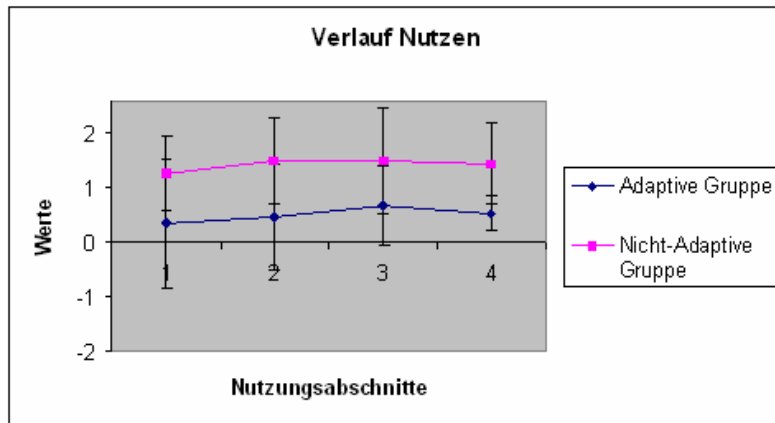
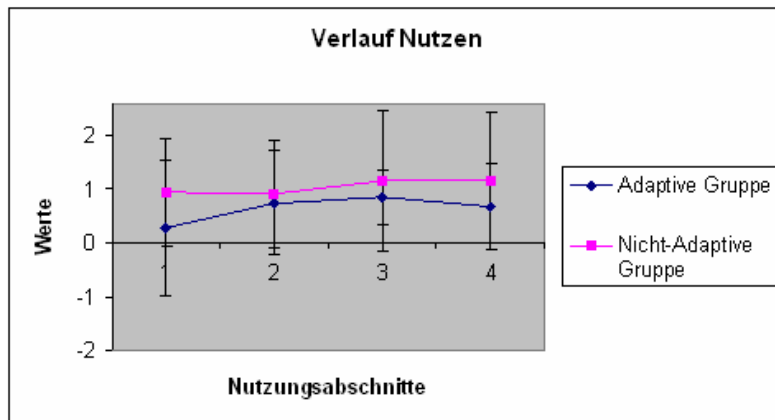


Abbildung 39 - Dritte Forschungsfrage: Ergebnis Kontrolle Fragebogen



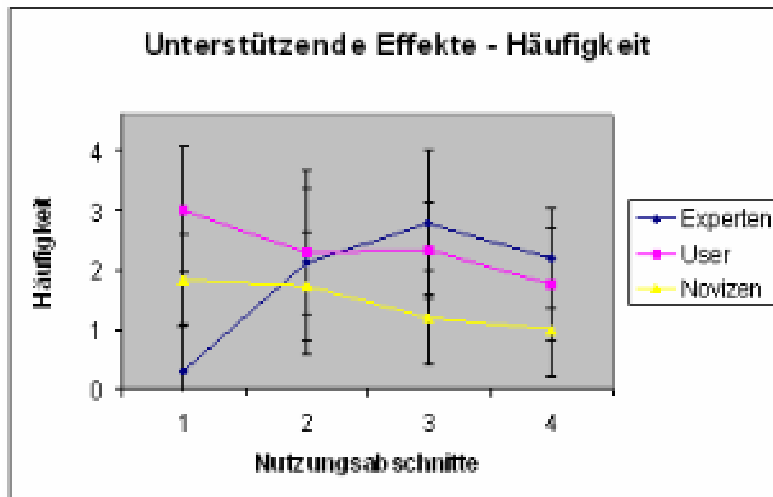
**Abbildung 40 - Dritte Forschungsfrage: Ergebnis Nutzen Tagebuch**



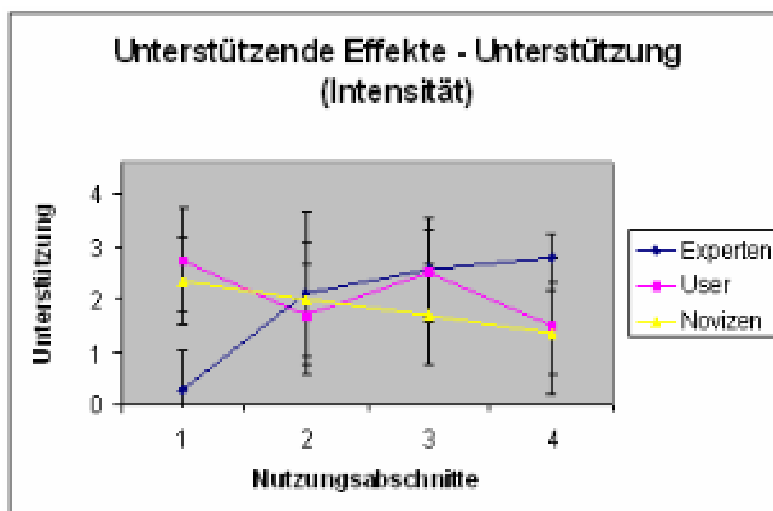
**Abbildung 41 - Dritte Forschungsfrage: Ergebnis Nutzen Fragebogen**



**Anhang L      Fünfte Forschungsfrage: Auszug signifikanter Ergebnisse**



**Abbildung 42 - Fünfte Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Unterstützende Effekte“ über den Studienzeitraum**



**Abbildung 43 - Fünfte Forschungsfrage: Ergebnis Parameter Intensität „Unterstützende Effekte“ über den Studienzeitraum**

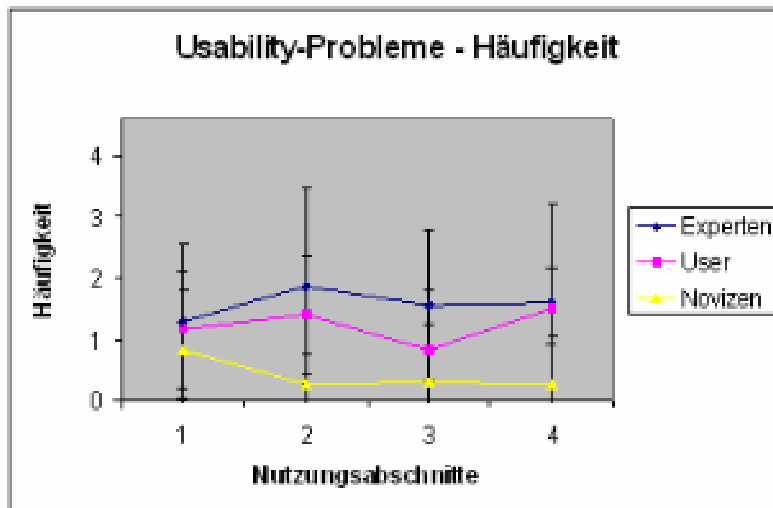


Abbildung 44 - Fünfte Forschungsfrage: Ergebnis Parameter geschätzte Häufigkeit „Usability-Probleme“ über den Studienzeitraum

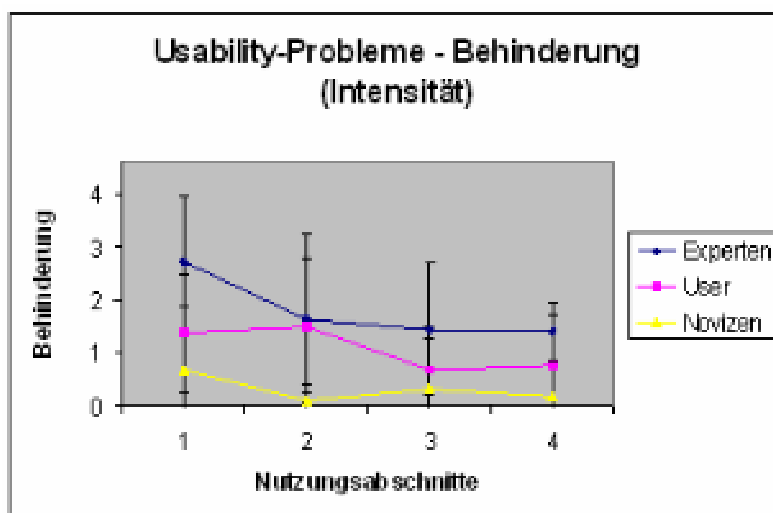


Abbildung 45 - Fünfte Forschungsfrage: Ergebnis Parameter Intensität „Usability-Probleme“ über den Studienzeitraum

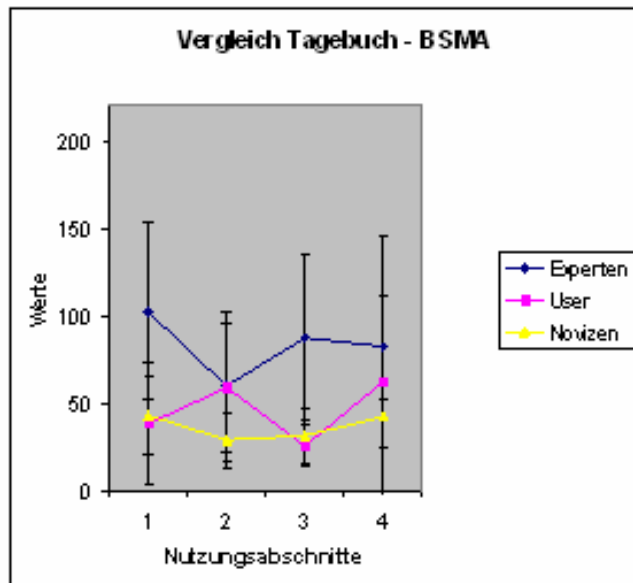


Abbildung 46 - Fünfte Forschungsfrage: Ergebnis BSMA Tagebuch

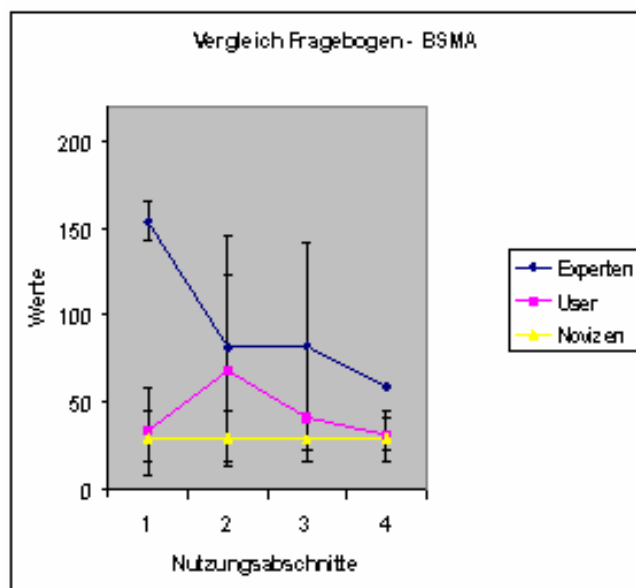
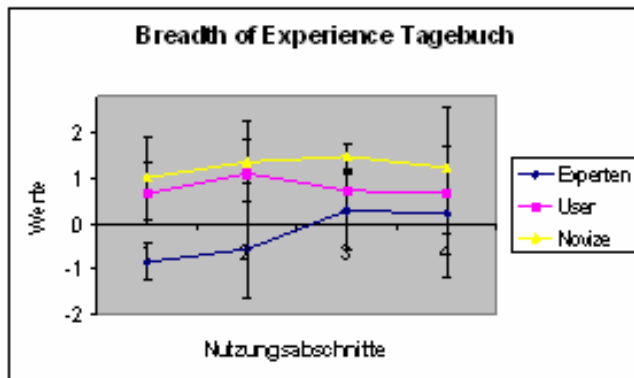
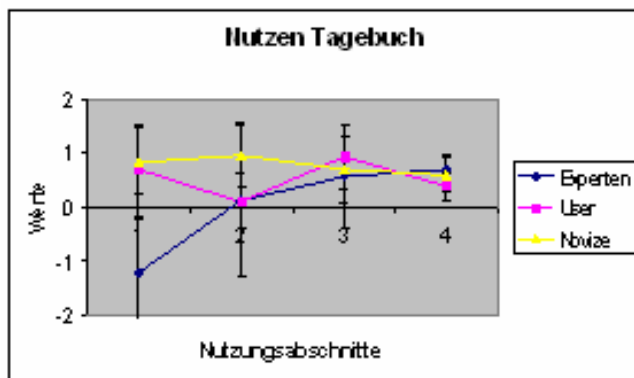


Abbildung 47 - Fünfte Forschungsfrage: Ergebnis BSMA Fragebogen



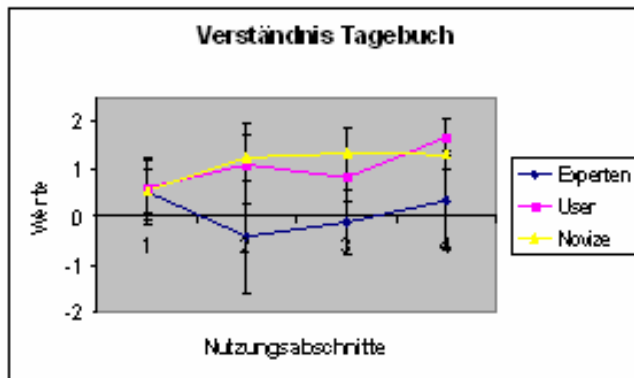
**Abbildung 48 - Fünfte Forschungsfrage: Ergebnis Breadth of Experience Tagebuch**

Signifikante Werte in den Nutzungsabschnitten 1 bis 3 von  $\alpha < 0,05$ , ermittelt durch den Mediantest für k unabhängige Stichproben.



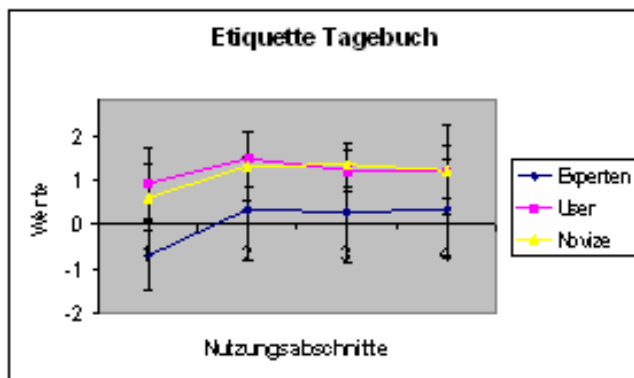
**Abbildung 49 - Fünfte Forschungsfrage: Ergebnis Nutzen Tagebuch**

Signifikante Werte in den Nutzungsabschnitten 1 und 2 von  $\alpha < 0,05$ , ermittelt durch den Mediantest für k unabhängige Stichproben.



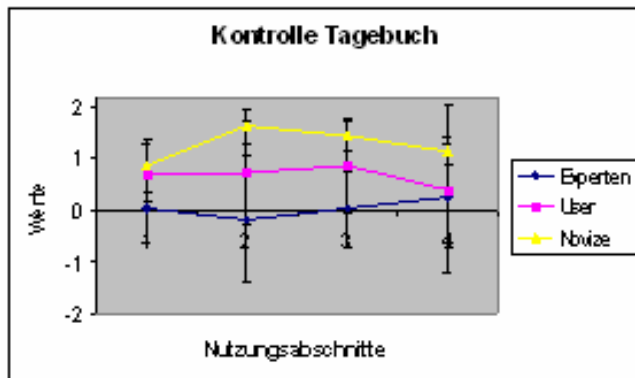
**Abbildung 50 - Fünfte Forschungsfrage: Ergebnis Verständnis Tagebuch**

Signifikante Werte in den Nutzungsabschnitten 2 und 3 von  $\alpha < 0,05$ , ermittelt durch den Mediantest für k unabhängige Stichproben.



**Abbildung 51 - Fünfte Forschungsfrage: Ergebnis Etiquette Tagebuch**

Signifikanter Wert in Nutzungsabschnitt 1 von  $\alpha < 0,05$ , ermittelt durch den Mediantest für k unabhängige Stichproben.



**Abbildung 52 - Fünfte Forschungsfrage: Ergebnis Kontrolle Tagebuch**

Signifikanter Wert in Nutzungsabschnitt 3 von  $\alpha < 0,05$ , ermittelt durch den Mediantest für k unabhängige Stichproben.